

On near optimal control of systems with slow observables

Vladimir Gaitsgory*

Sergey Rossomakhine[†]

August 12, 2016

Abstract

The paper deals with a problem of control of a system characterized by the fact that the influence of controls on the dynamics of certain functions of state variables (called observables) is relatively weak and the rates of change of these observables are much slower than the rates of change of the state variables themselves. The contributions of the paper are twofold. Firstly, the averaged system whose solutions approximate the trajectories of the slow observables is introduced, and it is shown that the optimal value of the problem of optimal control with time discounting criterion considered on the solutions of the system with slow observables (this problem is referred to as perturbed) converges to the optimal value of the corresponding problem of optimal control of the averaged system. Secondly, a way how an asymptotically optimal control of the perturbed problem can be constructed on the basis of an optimal solution of the averaged problem is indicated, sufficient and necessary optimality conditions for the averaged problem are stated, and a way how a near optimal solution of the latter can be constructed numerically is outlined (the construction being illustrated with an example).

Key words. Control of slow observables, Averaged system, Average control generating families, Occupational measures, Numerical solution

AMS subject classifications. 34E15, 34C29, 34A60, 93C70

1 Introduction and preliminaries

In this paper, we will study a problem of control of a system characterized by the fact that the influence of controls on the dynamics of certain functions of state variables (called *observables*) is relatively weak and the rates of change of these observables are much slower than the rates of change of the state variables themselves. Formally, a system with slow observables can be viewed as one embedded into a family of systems depending on a small parameter. When the parameter is set to zero, the observables remain constant. With non-zero but small positive values of the parameter, the observables may move slowly, and their dynamics can be influenced by controls.

In more detail, we will be considering the control system

$$\frac{dy_\epsilon(\tau)}{d\tau} = f(u(\tau), y_\epsilon(\tau)) + \epsilon g(u(\tau), y_\epsilon(\tau)), \quad y_\epsilon(0) = y_0, \quad (1)$$

*Department of Mathematics, Macquarie University, Sydney 2109, Australia, vladimir.gaitsgory@mq.edu.au; The work of V. Gaitsgory was supported by the Australian Research Council Discovery Grants DP130104432

[†]Flinders Mathematical Sciences Laboratory, School of Computer Science, Engineering and Mathematics, Flinders University, GPO Box 2100, Adelaide SA 5001, Australia, sergei.rossomakhine@flinders.edu.au; The work of S. Rossomakhine was supported by the Australian Research Council Discovery Grant DP150100618

where $\epsilon > 0$ is a small parameter; the vector functions $f(u, y) : U \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ and $g(u, y) : U \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ are continuous and satisfy Lipschitz conditions in y (uniformly in $u \in U$); the controls $u(\cdot)$ are measurable functions of time satisfying the inclusion $u(t) \in U$ (U being a given compact subset of a finite dimensional space). The system (1) will be referred to as the *perturbed system*.

Along with the perturbed system (1), we will be considering the *reduced system*

$$\frac{dy(\tau)}{d\tau} = f(u(\tau), y(\tau)), \quad (2)$$

(obtained from (1) by taking $\epsilon = 0$), and everywhere in what follows, it is assumed that

The reduced system (2) has k constants of motion or, more precisely, there exists a vector function $F(y) : \mathbb{R}^m \rightarrow \mathbb{R}^k$ such that

$$F'(y)f(u, y) = 0 \quad \forall u \in U, \quad \forall y \in Y, \quad (3)$$

where $F'(\cdot)$ is the Jacobian matrix and Y is a compact subset of \mathbb{R}^m in a neighbourhood of which the partial derivatives of the components of $F(\cdot)$ exist and satisfy Lipschitz conditions.

From (3) it follows that any solution $y(\tau)$ of the reduced system (2) that has the following property: if $F(y(0)) = z$, then $F(y(\tau)) = z \quad \forall \tau \geq 0$. That is, the set

$$Y_z \stackrel{\text{def}}{=} \{y \in \mathbb{R}^m : F(y) = z\} \quad (4)$$

is invariant with respect to the reduced system. (In what follows, it will be assumed that $Y_z \subset Y$ for z from a sufficiently large area.) This invariance property does not hold true for the solutions of the perturbed system (1). In fact, let $y_\epsilon(\tau)$ be the solution of (1) obtained with a control $u(\cdot)$ such that it is contained in Y for all $\tau \geq 0$. Then, taking the derivative of $z_\epsilon(\tau) \stackrel{\text{def}}{=} F(y_\epsilon(\tau))$, one arrives at the equation

$$\frac{dz_\epsilon(\tau)}{d\tau} = \epsilon h(u(\tau), y_\epsilon(\tau)), \quad z_\epsilon(0) = z_0 \stackrel{\text{def}}{=} F(y_0), \quad (5)$$

where

$$h(u, y) \stackrel{\text{def}}{=} F'(y)g(u, y). \quad (6)$$

(Below, we will be interested only in those solutions of (1) that are contained in Y for sufficiently small ϵ .) The components of the vector $z_\epsilon(\tau)$ are called *observables*. Note that the observables are changing their values slowly but these changes can be significant if the time period is long enough.

Systems with slow observables were studied mostly in the uncontrolled setting (see, e.g., [1], [7], [19], [31], [34], [41], [47]) with the only exception being problems of control of singularly perturbed Markov chains (see [52] and references therein). The latter present an important special case of problems that are dealt with in this paper. It is characterized by that $f(u, y) = A(u)y$ and $g(u, y) = B(u)y$, where $A(u), B(u)$ are matrix functions having some special properties (see Remark 2 below).

Note that the system consisting of (1) and (5) (for convenience, it will be referred to as the *augmented perturbed system*) is a special case of the system

$$\frac{dy_\epsilon(\tau)}{d\tau} = f_1(\epsilon, u(\tau), y_\epsilon(\tau)), \quad \frac{dz_\epsilon(\tau)}{d\tau} = \epsilon f_2(\epsilon, u(\tau), y_\epsilon(\tau), z_\epsilon(\tau)), \quad (7)$$

in which $f_1(\cdot)$ and $f_2(\cdot)$ are smooth in ϵ . Being rewritten in the “shrunk” time scale $t = \epsilon\tau$, such systems are commonly called weakly coupled singularly perturbed (**SP**) systems. Problems of control of SP systems (including weakly coupled SP systems) have received a great deal of attention (see [2], [3], [4], [5], [6], [11], [12], [13], [14], [17], [22], [23], [33], [35], [36], [39], [40], [43], [45], [46], [48], [50], [51] and references therein). However, to the best of our knowledge,

in all the earlier works, except the recent paper [45], the SP control systems were considered under the assumption that the set of occupational measures generated by the controls and the corresponding solutions of the reduced system

$$\frac{dy(\tau)}{d\tau} = f_1(0, u(\tau), y(\tau)) \quad (8)$$

is independent of the initial conditions from a sufficiently large set (the validity of this being guaranteed by imposing certain stability or controllability type conditions). In [45], a weakly coupled SP system similar to (7) was considered under the assumption that the reduced system (8) satisfies the non-expansivity condition introduced in [44], the latter does not exclude the situation when the set of occupational measures generated by the controls and the solutions of (8) depends on the initial conditions.

In our case, the set of occupational measures generated by the controls and the solutions of the reduced system (2) does depend on the initial conditions. In fact, due to the fact that the set Y_z is invariant with respect to the system, any measure generated by a control $u(\cdot)$ and the corresponding solution $y(\cdot)$ of (2) will be supported on $U \times Y_z$ if $y(0) \in Y_z$. That is, the sets of occupational measures corresponding to the initial conditions in Y_{z_1} and in Y_{z_2} ($z_1 \neq z_2$) are different. Also, the system (2) may not satisfy the non-expansivity condition of [44]. Thus, the available techniques are generally not applicable to the class of problems that we are considering.

The contributions of the paper are twofold. Firstly, the averaged system whose solutions approximate the trajectories of the slow observables is introduced, and it is shown that the optimal value of the optimal control problem

$$\epsilon \inf_{u(\cdot)} \int_0^{+\infty} e^{-(C\epsilon\tau)} r(u(\tau), y_\epsilon(\tau)) d\tau \stackrel{\text{def}}{=} R^*(\epsilon, y_0), \quad (9)$$

where $C > 0$ is a discount rate and \inf is sought over the controls and the corresponding solutions of the system (1), converges to the optimal value of an optimal control problem considered on the solutions of the averaged system. (These two problems will be referred to as *perturbed* and *averaged* problems respectively.) Note that the function $r(u, y) : U \times \mathbb{R}^m \mapsto \mathbb{R}$ is assumed to be continuous in (u, y) and satisfying Lipschitz conditions in y (uniformly with respect to $u \in U$) and that the multiplier ϵ before the integral in (9) is introduced for the values of the objective function to remain bounded with ϵ tending to zero.

Secondly, a way how an optimal or near optimal solution of the averaged problem can be found is described and a way how the latter can be used for construction of asymptotically optimal/near optimal solution of the perturbed problem is indicated. Note that the perturbed problem may become very ill-conditioned with small values of the parameter ϵ which makes it difficult (and sometimes even impossible) to tackle with the existing optimisation techniques. In the framework of the approach that we are proposing, a near optimal control of the perturbed problem is constructed on the basis of an optimal solution of the averaged problem that is independent of ϵ . Thus, the difficulties caused by the ill-conditioning are eliminated, and, in fact, the approximation of the optimal control improves as ϵ decreases towards 0.

The paper consists of seven sections and is organized as follows. Section 1 is this introduction. In Section 2, we introduce the averaged system and the averaged optimal control problem, and we establish that, under the validity of certain assumptions, these can be used for an approximation of the perturbed system and the perturbed optimal control problem. In Section 3, the general assumptions used in Section 2 are shown to be valid if certain stability or controllability conditions are satisfied on the invariant sets of the reduced system. In Section 4, the assumptions of Section 2 are verified to be true for a special case when the reduced system is uncontrolled. In Section 5, we give a definition of an average control generating (ACG) family and establish sufficient and necessary conditions for an ACG family to be optimal in the averaged problem, and we also outline an algorithm for finding near optimal ACG families (based on results of [30]). In Section 6, we discuss a way how asymptotically optimal/near optimal controls of the

perturbed problem can be constructed on the basis of optimal/near optimal ACG families, and we illustrate the construction with a numerical example. In Section 7, we give proofs of results stated in Sections 2 and 6.

Let us conclude this section with some notations and definitions. Given a compact metric space W , $\mathcal{B}(W)$ will stand for the σ -algebra of its Borel subsets and $\mathcal{P}(W)$ will denote the set of probability measures defined on $\mathcal{B}(W)$. The set $\mathcal{P}(W)$ will always be treated as a compact metric space with a metric ρ , which is consistent with its weak* topology. That is, a sequence $\gamma^k \in \mathcal{P}(W)$, $k = 1, 2, \dots$, converges to $\gamma \in \mathcal{P}(W)$ in this metric if and only if

$$\lim_{k \rightarrow \infty} \int_W \phi(w) \gamma^k(dw) = \int_W \phi(w) \gamma(dw),$$

for any continuous $\phi(w) : W \rightarrow \mathbb{R}^1$.

Using this metric ρ , one can define the Hausdorff metric ρ_H on the set of subsets of $\mathcal{P}(W)$ as follows: $\forall \Gamma_i \subset \mathcal{P}(W)$, $i = 1, 2$,

$$\rho_H(\Gamma_1, \Gamma_2) \stackrel{\text{def}}{=} \max\left\{\sup_{\gamma \in \Gamma_1} \rho(\gamma, \Gamma_2), \sup_{\gamma \in \Gamma_2} \rho(\gamma, \Gamma_1)\right\}, \quad (10)$$

where $\rho(\gamma, \Gamma_i) \stackrel{\text{def}}{=} \inf_{\gamma' \in \Gamma_i} \rho(\gamma, \gamma')$. Note that, although, by some abuse of terminology, we refer to $\rho_H(\cdot, \cdot)$ as to a metric on the set of subsets of $\mathcal{P}(Y \times U)$, it is, in fact, a semi metric on this set (note that $\rho_H(\Gamma_1, \Gamma_2) = 0$ implies $\Gamma_1 = \Gamma_2$ if Γ_1 and Γ_2 are closed but the equality may not be true if at least one of these sets is not closed).

Given a measurable function $w(\cdot) : [0, \infty) \rightarrow W$, the *occupational measure* generated by this function on the interval $[0, S]$ is the probability measure $\gamma^{w(\cdot), S} \in \mathcal{P}(W)$ defined by the equation

$$\gamma^{w(\cdot), S}(Q) \stackrel{\text{def}}{=} \frac{1}{S} \int_0^S 1_Q(w(t)) dt, \quad \forall Q \in \mathcal{B}(W), \quad (11)$$

where $1_Q(\cdot)$ is the indicator function. The occupational measure generated by this function on the interval $[0, \infty)$ is the probability measure $\gamma^{w(\cdot)} \in \mathcal{P}(W)$ defined as the limit (assumed to exist)

$$\gamma^{w(\cdot)}(Q) \stackrel{\text{def}}{=} \lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S 1_Q(w(t)) dt, \quad \forall Q \in \mathcal{B}(W). \quad (12)$$

Note that (11) is equivalent to that

$$\int_W q(w) \gamma^{w(\cdot), S}(dw) = \frac{1}{S} \int_0^S q(w(t)) dt \quad (13)$$

for any $q(\cdot) \in C(W)$, and (12) is equivalent to that

$$\int_W q(w) \gamma^{w(\cdot)}(dw) = \lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S q(w(t)) dt \quad (14)$$

for any $q(\cdot) \in C(W)$. Note that, if $w(\cdot)$ is T -periodic, then (14) becomes

$$\int_W q(w) \gamma^{w(\cdot)}(dw) = \frac{1}{T} \int_0^T q(w(t)) dt. \quad (15)$$

2 Averaged system and averaged optimal control problem

It has been understood that, in dealing with uncontrolled systems with slow observables ($f(u, y) = f(y)$, $g(u, y) = g(y)$), the trajectories of the slow observables can be approximated by solutions of the averaged system (see, e.g., Theorem 6.5 in [7])

$$\frac{dz(\tau)}{d\tau} = \epsilon \tilde{h}(\mu_{z(\tau)}), \quad (1)$$

where $\tilde{h}(\mu_z) \stackrel{\text{def}}{=} \int_{Y_z} F'(y)g(y)\mu_z(dy)$ and where $\mu_z(dy)$ is the occupational measure generated (induced) by the solutions of the reduced system that are contained in Y_z .

In the presence of controls, there are multiple occupational measures that are generated by the controls and the corresponding solutions of the reduced system staying in Y_z . The closed convex hull of the set of such measures is presentable in the form (see Theorem 2.1 in [23])

$$W(z) \stackrel{\text{def}}{=} \{\mu \in \mathcal{P}(U \times Y) : \int_{U \times Y_z} \nabla \phi(y)^T f(u, y) \mu(du, dy) = 0 \quad \forall \phi(\cdot) \in C^1, \quad \text{supp}(\mu) \subset U \times Y_z\}.^1 \quad (2)$$

It is, therefore, natural to consider the following averaged system

$$\frac{dz(\tau)}{d\tau} = \epsilon \tilde{h}(\mu(\tau)), \quad z(0) = z_0, \quad (3)$$

where

$$\tilde{h}(\mu) \stackrel{\text{def}}{=} \int_{U \times Y} F'(y)g(u, y)\mu(du, dy) \quad (4)$$

and where

$$\mu(\tau) \in W(z(\tau)). \quad (5)$$

In contrast to (1), the system (3) is controlled, with the role of controls being played by the measure-valued functions $\mu(\tau)$ satisfying the inclusion (5). Note that the map $W(z) : Z \rightsquigarrow \mathcal{P}(U \times Y)$ defined in (2) is convex and weak* compact-valued.

By changing the time scale $t = \epsilon\tau$, one can rewrite (3), (5) in the form

$$\frac{dz(t)}{dt} = \tilde{h}(\mu(t)), \quad z(0) = z_0, \quad (6)$$

$$\mu(t) \in W(z(t)). \quad (7)$$

In what follows, it is the system (6) that will be referred to as the *averaged system*. A pair $(\mu(\cdot), z(\cdot))$ that satisfies (6) and (7) will be called a *solution of the averaged system*.

Consider the following optimal control problem

$$\inf_{(\mu(\cdot), z(\cdot))} \int_0^{+\infty} e^{-Ct} \tilde{r}(\mu(t)) dt \stackrel{\text{def}}{=} \tilde{R}^*(y_0), \quad (8)$$

where

$$\tilde{r}(\mu) \stackrel{\text{def}}{=} \int_{U \times Y} r(u, y) \mu(du, dy) \quad (9)$$

and *inf* is sought over all solutions of the averaged system. This will be referred to as the *averaged optimal control problem*.

Note that in the shrank time scale $t = \tau\epsilon$, the problem (9) is rewritten in the form

$$\inf_{u(\cdot)} \int_0^{+\infty} e^{-Ct} r(u(t), y_\epsilon(t)) dt \stackrel{\text{def}}{=} R^*(\epsilon, y_0), \quad (10)$$

where *inf* is over the controls and the corresponding solutions of the system

$$\frac{dy_\epsilon(t)}{dt} = \frac{1}{\epsilon} f(u(t), y_\epsilon(t)) + g(u(t), y_\epsilon(t)), \quad y(0) = y_0. \quad (11)$$

This is the shrank time scale version of the perturbed system (1). Note that in this time scale, the system (5) describing the dynamics of the observables takes the form

$$\frac{dz_\epsilon(t)}{dt} = h(u(t), y_\epsilon(t)), \quad z_\epsilon(0) = z_0. \quad (12)$$

¹Given a compact set X and $\mu \in \mathcal{P}(X)$, $\text{supp}(\mu)$ stands for the support of μ , that is, the “minimal” closed subset of X the μ -measure of which is equal to 1.

In this section, we will present results establishing that trajectories of the observables are approximated by solutions of the averaged system and that the optimal value of the problem (10) is approximated by the optimal value of the averaged optimal control problem (9). These results are valid under certain assumptions that are introduced below.

Let Z be a compact subset of \mathbb{R}^k such that $Y_z \subset Y$ for any $z \in Z$. In stating our results, we will be assuming (without mentioning it) that the multivalued map $Y_z : Z \rightsquigarrow Y$ is Lipschitz continuous on Z . That is,

$$d_H(Y_{z'}, Y_{z''}) \leq L_0 \|z' - z''\| \quad \forall z', z'' \in Z, \quad L_0 = \text{const.} \quad (13)$$

(Here and in what follows, $d_H(\cdot, \cdot)$ stands for the Hausdorff metric defined on compact subsets of a finite-dimensional space.) Note that the Lipschitz continuity of Y_z can be verified to be true if there exists a $m \times (m - k)$ matrix D (which may depend on y) such that the inverse of the matrix $\hat{D}(y) \stackrel{\text{def}}{=} (F'(y)^T, D)$ exists and is uniformly bounded: $\|\hat{D}(y)^{-1}\| \leq c = \text{const} \quad \forall y \in Y$ (see examples in Sections 3 and 4).

Assumption 1 *The multivalued map $W(z) : Z \rightsquigarrow \mathcal{P}(U \times Y)$ is continuous and the multivalued map $V(z) : Z \rightsquigarrow \mathbb{R}^{k+1}$,*

$$V(z) \stackrel{\text{def}}{=} \cup_{\mu \in W(z)} \{(\tilde{h}(\mu), \tilde{r}(\mu))\}, \quad (14)$$

is Lipschitz continuous. That is,

$$d_H(V(z'), V(z'')) \leq L \|z' - z''\| \quad \forall z', z'' \in Z, \quad L = \text{const.} \quad (15)$$

Sufficient conditions for Assumption 1 (and for Assumption 2 introduced below) to be satisfied are discussed in Sections 3 and 4.

The following proposition is important for understanding of the approximation results stated below.

Proposition 1 *If the multi-valued map $W(\cdot) : Z \rightsquigarrow \mathcal{P}(U \times Y)$ is continuous, then the occupational measure $\mu^{u(\cdot), y(\cdot), S}$ generated on the interval $[0, S]$ by an arbitrary control $u(\cdot)$ and the solution $y(\cdot)$ of the system (2) obtained with this control and an initial condition $y(0) \in Y_z$ satisfies the inequality*

$$\rho(\mu^{u(\cdot), y(\cdot), S}, W(z)) \leq \beta(S), \quad \lim_{S \rightarrow \infty} \beta(S) = 0, \quad (16)$$

the estimate being uniform with respect to $y(0) \in Y_z$ and $z \in Z$.

Proof. The proof of the proposition is given at the end of this section. \square

Denote by $\mathcal{M}_z(S, y(0))$ the set of occupational measures generated by all pairs $(u(\cdot), y(\cdot))$ on an interval $[0, S]$, where $u(\cdot)$ is a control and $y(\cdot)$ is the corresponding solution of the reduced system (2) such that $y(0) \in Y_z$. Then the estimate (16) of Proposition 1 can be rewritten in the form

$$\sup_{\mu \in \mathcal{M}_z(S, y(0))} \rho(\mu, W(z)) \leq \beta(S) \quad \forall y(0) \in Y_z, \quad \forall z \in Z. \quad (17)$$

Assumption 2 *The following estimate is valid*

$$\max_{\mu \in W(z)} \rho(\mu, \mathcal{M}_z(S, y(0))) \leq \beta(S) \quad \forall y(0) \in Y_z, \quad \forall z \in Z, \quad \lim_{S \rightarrow \infty} \beta(S) = 0. \quad (18)$$

Note that, by the definition of the Hausdorff metric (see (10)), the validity of (17) and (18) is equivalent to

$$\rho_H(\mathcal{M}_z(S, y(0)), W(z)) \leq \beta(S) \quad \forall y(0) \in Y_z, \quad \forall z \in Z, \quad (19)$$

where $\lim_{S \rightarrow \infty} \beta(S) = 0$. The validity of (19) and, hence, the validity of Assumption 2 can be verified with the help of the following statement (see instances of application of the latter in Sections 3 and 4).

Proposition 2 *The estimate (19) is valid if and only if, for any pair $(u^1(\cdot), y^1(\cdot))$, where $u^1(\cdot)$ is an arbitrary control and $y^1(\cdot)$ is the solution of (2) obtained with this control and $y^1(0) = y_0^1 \in Y_z$ and for any $y_0^2 \in Y_z$, there exists a control $u^2(\cdot)$ such that, for an arbitrary Lipschitz continuous function $q(u, y)$,*

$$\left| \frac{1}{S} \int_0^S q(u^1(\tau), y^1(\tau)) d\tau - \frac{1}{S} \int_0^S q(u^2(\tau), y^2(\tau)) d\tau \right| \leq \beta_q(S), \quad \lim_{S \rightarrow \infty} \beta_q(S) = 0 \quad (20)$$

where $y^2(\cdot)$ is the solution of the system (2) obtained with the control $u^2(\cdot)$ and with the initial condition $y^2(0) = y_0^2$.

Proof. The proof follows from Theorem 2.1(iii) and Proposition 4.1 in [23]. \square

The main results of this section are the following two theorems.

Theorem 1 *Let Assumption 1 be satisfied and let the averaged system be viable in Z (see [8]). Then, corresponding to any pair $(u_\epsilon(t), y_\epsilon(t))$, where $u_\epsilon(t)$ is a control and $y_\epsilon(t)$ is the solution of (11) obtained with this control such that*

$$y_\epsilon(t) \in Y, \quad F(y_\epsilon(t)) \in \check{Z} \subset \text{int}Z \quad \forall t \in [0, \infty) \quad (21)$$

(\check{Z} being a compact subset of the interior of Z), there exists a solution $(\mu(t), z(t))$ of the averaged system (6), (7) such that $z(t) \in Z \quad \forall t \in [0, \infty)$ and such that

$$\max_{t \in [0, T]} \|F(y_\epsilon(t)) - z(t)\| \leq \beta(\epsilon, T), \quad (22)$$

$$\left| \int_0^T e^{-Ct} r(u_\epsilon(t), y_\epsilon(t)) dt - \int_0^T e^{-Ct} \tilde{r}(\mu(t)) dt \right| \leq \beta(\epsilon, T), \quad (23)$$

where $\lim_{\epsilon \rightarrow 0} \beta(\epsilon, T) = 0$.

Theorem 2 *Let Assumptions 1 and 2 be valid, and let the augmented perturbed system (11), (12) be viable in $Y \times Z$. Then, corresponding to any solution $(\mu(t), z(t))$ of the averaged system (6), (7) such that*

$$z(t) \in \check{Z} \subset \text{int}Z, \quad Y_{z(t)} \subset \check{Y} \subset \text{int}Y \quad \forall t \in [0, \infty) \quad (24)$$

(\check{Z} and \check{Y} being compact subsets of the interiors of Y and Z respectively), there exists a control $u_\epsilon(t)$ such that the corresponding solution $y_\epsilon(t)$ of the system (11) satisfies the inclusions

$$(y_\epsilon(t), F(y_\epsilon(t))) \in Y \times Z \quad \forall t \in [0, \infty), \quad (25)$$

and the estimates (22) and (23) are valid.

Proof. The proofs of the theorems are given in Section 7. \square

Corollary 1 *Let Assumption 1 be satisfied and the averaged system be viable in Z . Let also there exist a control $u_\epsilon(\tau)$ such that the corresponding solution $y_\epsilon(\tau)$ of the system (11) satisfies the inclusions (21) and also*

$$\liminf_{\epsilon \rightarrow 0} \int_0^{+\infty} e^{-Ct} r(u_\epsilon(t)(t), y_\epsilon(t)) dt = \liminf_{\epsilon \rightarrow 0} R^*(\epsilon, y_0). \quad (26)$$

Then the optimal values of the problems (10) and (8) are related by the inequality

$$\liminf_{\epsilon \rightarrow 0} R^*(\epsilon, y_0) \geq \tilde{R}^*(y_0). \quad (27)$$

Proof. Let us choose a function $T(\epsilon)$ in such a way that

$$\lim_{\epsilon \rightarrow 0} T(\epsilon) = \infty, \quad \lim_{\epsilon \rightarrow 0} \beta(\epsilon, T(\epsilon)) = 0 \quad (28)$$

(the fact that such $T(\epsilon)$ exists is readily verifiable). By Theorem 1, there exists a solution $(\mu(t), z(t))$ of the averaged system (6) such that the estimate (23) is valid. Then

$$\left| \int_0^\infty e^{-Ct} r(u_\epsilon(t), y_\epsilon(t)) dt - \int_0^\infty e^{-Ct} \tilde{r}(\mu(t)) dt \right| \leq \frac{2}{C} \max_{(u,y) \in U \times Y} |r(u,y)| e^{-CT(\epsilon)} + \beta(\epsilon, T(\epsilon)) \stackrel{\text{def}}{=} \bar{\beta}(\epsilon), \quad (29)$$

where $\lim_{\epsilon \rightarrow 0} \bar{\beta}(\epsilon) = 0$ (due to (28)). Consequently,

$$\int_0^\infty e^{-Ct} r(u_\epsilon(t), y_\epsilon(t)) dt \geq \int_0^\infty e^{-Ct} \tilde{r}(\mu(t)) dt - \bar{\beta}(\epsilon) \geq \tilde{R}^*(y_0) - \bar{\beta}(\epsilon).$$

By (26), the latter implies (27). \square

Corollary 2 *Let Assumptions 1 and 2 be satisfied. Let the averaged system be viable in Z and the augmented perturbed system be viable in $Y \times Z$. Let also there exist a control $u_\epsilon(\tau)$ such that the corresponding solution $y_\epsilon(\tau)$ of (11) satisfies (21), (26) and there exist an optimal solution $(\mu(t), z(t))$ of the averaged problem that satisfies (24). Then the limit of the optimal value of the problem (10) exists and is equal to the optimal value of the averaged problem (8):*

$$\lim_{\epsilon \rightarrow 0} R^*(\epsilon, y_0) = \tilde{R}^*(y_0). \quad (30)$$

Proof. Similarly to the proof of Corollary 1, it is established that from Theorem 2 it follows that

$$\limsup_{\epsilon \rightarrow 0} R^*(\epsilon, y_0) \leq \tilde{R}^*(y_0). \quad (31)$$

This along with (27) prove (30). \square

Remark 1 The statements of the Theorems 1 and 2 are simplified if one assumes that Z is invariant with respect to the solutions of the averaged system and that $Y \times Z$ is invariant with respect to the solutions of the augmented perturbed system. In this case, corresponding to any pair $(u_\epsilon(t), y_\epsilon(t))$, where $u_\epsilon(t)$ is an arbitrary control and $y_\epsilon(t)$ is the solution of (11) obtained with this control, there exists a solution $(\mu(t), z(t))$ of the averaged system such that (22), (23) are satisfied (provided that Assumption 1 is valid). Conversely, corresponding to any solution $(\mu(t), z(t))$ of the averaged system, there exists a control $u_\epsilon(t)$ such that (22), (23) are satisfied with $y_\epsilon(t)$ being the solution of (11) obtained with the control $u_\epsilon(t)$ (provided that Assumptions 1 and 2 are valid). Also (30) is valid in this case as well.

Proof of Proposition 1. Assume the statement of the proposition is not true. Then there exist a number $\alpha > 0$ and sequences $S_i, z_i \in Z, u^i(\cdot), y^i(\cdot)$ ($y^i(0) \in Y_{z_i}$), $i = 1, 2, \dots$, such that

$$\rho(\mu^{u^i(\cdot), y^i(\cdot), S_i}, W(z_i)) \geq \alpha, \quad \lim_{i \rightarrow \infty} S_i = \infty. \quad (32)$$

Due to compactness of Z and due to compactness of $\mathcal{P}(U \times Y)$, one may assume (without loss of generality) that there exist limits

$$\lim_{i \rightarrow \infty} z_i = z \in Z, \quad \lim_{i \rightarrow \infty} \mu^{u^i(\cdot), y^i(\cdot), S_i} = \mu \in \mathcal{P}(U \times Y). \quad (33)$$

Also, $\lim_{i \rightarrow \infty} W(z_i) = W(z)$ (due to assumed continuity of $W(\cdot)$). Hence, from (32) it follows that

$$\rho(\mu, W(z)) \geq \alpha. \quad (34)$$

By the definition of the occupational measure (see (13)), for any $\phi(\cdot) \in C^1$,

$$\begin{aligned} \lim_{i \rightarrow \infty} \int_{U \times Y} \nabla \phi(y)^T f(u, y) \mu^{u^i(\cdot), y^i(\cdot), S_i}(du, dy) &= \lim_{i \rightarrow \infty} \frac{1}{S_i} \int_0^{S_i} \phi(y^i(\tau)) f(u^i(\tau), y^i(\tau)) \\ &= \lim_{i \rightarrow \infty} \frac{1}{S_i} [\phi(y^i(S_i)) - \phi(y^i(0))] = 0. \end{aligned}$$

Consequently, due to the second equality in (33),

$$\int_{U \times Y} \nabla \phi(y)^T f(u, y) \mu(du, dy) = 0. \quad (35)$$

From the fact that Y_z is upper semicontinuous it follows that, for any $\kappa > 0$, $Y_{z_i} \subset Y_z + \kappa \bar{B}$ if i is large enough (\bar{B} being the closed unit ball in \mathbb{R}^m). Consequently,

$$1 = \mu^{u^i(\cdot), y^i(\cdot), S_i}(U \times Y_{z_i}) = \mu^{u^i(\cdot), y^i(\cdot), S_i}(U \times (Y_z + \kappa \bar{B})),$$

and

$$\mu(U \times (Y_z + \kappa \bar{B})) = \lim_{i \rightarrow \infty} \mu^{u^i(\cdot), y^i(\cdot), S_i}(U \times (Y_z + \kappa \bar{B})) = 1.$$

Since κ is arbitrary small, it follows that

$$\mu(U \times Y_z) = 1 \Rightarrow \text{supp}(\mu) \subset U \times Y_z.$$

The latter along with (35) imply that $\mu \in W(z)$, which contradicts (34). \square

3 Verification of Assumptions 1 and 2

In this section we will discuss some stability and controllability type conditions that ensure the validity of Assumptions 1 and 2.

Proposition 3 *Let there exist a positive definite matrix Q and a positive number α such that*

$$(f(u, y') - f(u, y''))^T Q(y' - y'') \leq -\alpha \|y' - y''\|^2 \quad \forall u \in U, \quad \forall y', y'' \in Y_z, \quad z \in Z. \quad (1)$$

Then Assumptions 1 and 2 are satisfied.

Proof. Let us first establish that Assumption 2 is satisfied. By Proposition 2, to prove the validity of the latter, it is sufficient to show that, given a pair $(u^1(\cdot), y^1(\cdot))$, where $u^1(\cdot)$ is an arbitrary control and $y^1(\cdot)$ is a solution of (2) obtained with this control and an initial condition $y^1(0) = y_0^1 \in Y_z$ and given $y_0^2 \in Y_z$, there exists a control $u^2(\cdot)$ such that, for an arbitrary Lipschitz continuous function $q(u, y)$, the estimate (20) is valid (with $y^2(\cdot)$ being the solution of the system (2) obtained with the control $u^2(\cdot)$ and with the initial condition $y^2(0) = y_0^2$).

Take $u^2(\cdot)$ to be equal to $u^1(\cdot)$. Then, by (1),

$$\begin{aligned} \frac{d}{d\tau} (y^1(\tau) - y^2(\tau))^T Q (y^1(\tau) - y^2(\tau)) &= 2 (f(u^1(\tau), y^1(\tau)) - f(u^1(\tau), y^2(\tau)))^T Q (y^1(\tau) - y^2(\tau)) \\ &\leq -2\alpha \|y^1(\tau) - y^2(\tau)\|^2 \leq -\frac{2\alpha}{\lambda_{\max}} (y^1(\tau) - y^2(\tau))^T Q (y^1(\tau) - y^2(\tau)), \end{aligned}$$

where λ_{\max} is the maximal eigenvalue of Q . It follows that

$$(y^1(\tau) - y^2(\tau))^T Q (y^1(\tau) - y^2(\tau)) \leq e^{-\frac{2\alpha}{\lambda_{\max}} \tau} (y_0^1 - y_0^2)^T Q (y_0^1 - y_0^2) \quad \forall \tau \geq 0,$$

the latter implying

$$\|y^1(\tau) - y^2(\tau)\| \leq \beta_1 e^{-\beta_2 \tau} \|y_0^1 - y_0^2\| \quad \forall \tau \geq 0, \quad (2)$$

where β_1, β_2 are constants ($\beta_1 > 1, \beta_2 > 0$). Using (2) and Lipschitz continuity of $q(\cdot)$, one obtains

$$\begin{aligned} \left| \frac{1}{S} \int_0^S q(u^1(\tau), y^1(\tau)) d\tau - \frac{1}{S} \int_0^S q(u^1(\tau), y^2(\tau)) d\tau \right| &\leq L_q \frac{1}{S} \left(\int_0^S \|y^1(\tau) - y^2(\tau)\| d\tau \right) \\ &\leq L_q \beta_1 \|y_0^1 - y_0^2\| \left(\frac{1}{S} \int_0^S e^{-\beta_2 \tau} d\tau \right) \leq \frac{1}{S} \left(L_q \frac{\beta_1}{\beta_2} \max_{y', y'' \in Y} \|y' - y''\| \right), \end{aligned}$$

where L_q is a Lipschitz constant of $q(\cdot)$. Thus, (20) is satisfied with $\beta_q(S) = O(\frac{1}{S})$, and the validity of Assumption 2 is established.

To prove the validity of Assumption 1, it is sufficient to show that, for any $z', z'' \in Z$,

$$\min_{\mu'' \in W(z'')} \left| \int_{U \times Y_{z'}} q(u, y) \mu'(du, dy) - \int_{U \times Y_{z''}} q(u, y) \mu''(du, dy) \right| \leq \hat{\kappa}_q(\|z' - z''\|) \quad \forall \mu' \in W(z'), \quad (3)$$

where $\lim_{\theta \rightarrow 0} \hat{\kappa}_q(\theta) = 0$ for any continuous $q(u, y)$, and that

$$\min_{\mu'' \in W(z'')} \left| \int_{U \times Y_{z'}} q(u, y) \mu'(du, dy) - \int_{U \times Y_{z''}} q(u, y) \mu''(du, dy) \right| \leq \hat{L}_q \|z' - z''\| \quad \forall \mu' \in W(z'), \quad (4)$$

where \hat{L}_q is a positive constant if $q(u, y)$ satisfies Lipschitz conditions in y (uniformly in $u \in U$).

Let $\mu' \in W(z')$. By (18), there exist a control $u'(\cdot)$ and the corresponding solution $y'(\cdot)$ of the system (2) satisfying an initial condition $y'(0) = y'_0 \in Y_{z'}$ such that

$$\left| \int_{U \times Y_{z'}} q(u, y) \mu'(du, dy) - \frac{1}{S} \int_0^S q(u'(\tau), y'(\tau)) d\tau \right| \leq \beta_q(S) \quad \text{with} \quad \lim_{S \rightarrow \infty} \beta_q(S) = 0. \quad (5)$$

Let $y''_0 \stackrel{\text{def}}{=} \operatorname{argmin}_{y \in Y_{z''}} \{\|y - y'_0\|\}$. By (13),

$$\|y'_0 - y''_0\| \leq L_0 \|z' - z''\|. \quad (6)$$

Denote by $y''(\cdot)$ the solution of the system (2) obtained with the control $u'(\cdot)$ that satisfies the initial condition $y''(0) = y''_0$ (note that $y''(\tau) \in Y_{z''} \quad \forall \tau \geq 0$). It will be shown below that

$$\|y'(\tau) - y''(\tau)\| \leq \hat{L} \|z' - z''\| \quad \forall \tau \geq 0, \quad (7)$$

where \hat{L} is a positive constant. Using this inequality, one can obtain

$$\left| \frac{1}{S} \int_0^S q(u'(\tau), y'(\tau)) d\tau - \frac{1}{S} \int_0^S q(u'(\tau), y''(\tau)) d\tau \right| \leq \kappa_q(\hat{L} \|z' - z''\|), \quad (8)$$

where

$$\kappa_q(\theta) \stackrel{\text{def}}{=} \max_{u \in U} \max_{y', y'' \in Y} \{|q(u, y') - q(u, y'')| \mid \|y' - y''\| \leq \theta\}.$$

(Note that $\lim_{\theta \rightarrow 0} \kappa_q(\theta) = 0$.) Denote by $\mu''_S \in \mathcal{M}_{z''}(S, y''_0)$, the occupational measure generated by the pair $(u'(\cdot), y''(\cdot))$ on the interval $[0, S]$. From the validity of (19) it follows that there exists $\bar{\mu}'' \in W(z'')$ such that

$$\rho(\mu''_S, \bar{\mu}'') \leq \beta(S), \quad \lim_{S \rightarrow \infty} \beta(S) = 0,$$

which implies

$$\left| \frac{1}{S} \int_0^S q(u'(\tau), y''(\tau)) d\tau - \int_{U \times Y_{z''}} q(u, y) \bar{\mu}''(du, dy) \right| \leq \bar{\beta}_q(S), \quad \lim_{S \rightarrow \infty} \bar{\beta}_q(S) = 0. \quad (9)$$

The latter along with (5) and (8) imply

$$|\int_{U \times Y_{z'}} q(u, y) \mu'(du, dy) - \int_{U \times Y_{z''}} q(u, y) \bar{\mu}''(du, dy)| \leq \kappa_q(\hat{L} \|z' - z''\|) + \beta_q(S) + \bar{\beta}_q(S),$$

which lead to

$$\min_{\mu'' \in W(z'')} |\int_{U \times Y_{z'}} q(u, y) \mu'(du, dy) - \int_{U \times Y_{z''}} q(u, y) \mu''(du, dy)| \leq \kappa_q(\hat{L} \|z' - z''\|) + \beta_q(S) + \bar{\beta}_q(S).$$

Passing to the limit as $S \rightarrow \infty$ in the expression above, one obtains (3) with $\hat{\kappa}_q(\theta) = \kappa_q(\hat{L}\theta)$. It also leads to (4) with $\hat{L}_q = \hat{L}L_q$ for a function $q(\cdot)$ that is Lipschitz continuous in y with a constant L_q (since $\kappa_q(\theta) \leq L_q\theta$ in this case).

To finalize the proof of the proposition, one now needs to establish the validity of (7). Let $\sigma \in (0, 1)$ be such that

$$\beta_1 e^{-\beta_2 \sigma} = \frac{1}{2} \quad (10)$$

(that is, $\sigma = \frac{1}{\beta_2} \ln 2\beta_1$) and let $\tau_l \stackrel{\text{def}}{=} l\sigma$, $l = 0, 1, \dots$. Define y_l'' by the equation $y_l'' \stackrel{\text{def}}{=} \operatorname{argmin}_{y \in Y_{z''}} \{\|y - y'(\tau_l)\|\}$. Similarly to (6), from (13) it follows that

$$\|y'(\tau_l) - y_l''\| \leq L_0 \|z' - z''\|, \quad l = 0, 1, \dots \quad (11)$$

Let $y_l''(\tau)$ stand for the solution of the system (2) considered on the interval $[\tau_l, \tau_{l+1}]$ that is obtained with the control $u'(\tau)$ and that satisfies the initial condition $y_l''(\tau_l) = y_l''$ (note that $y_l''(\tau) \in Y_{z''}$). Using (11) and Gronwall- Bellman lemma, one can readily verify that

$$\max_{\tau \in [\tau_l, \tau_{l+1}]} \|y'(\tau) - y_l''(\tau)\| \leq \hat{L}_0 \|z' - z''\|, \quad l = 0, 1, \dots, \quad (12)$$

where \hat{L}_0 is an appropriately chosen constant. Also, similarly to (2), one can obtain that

$$\|y''(\tau) - y_l''(\tau)\| \leq \beta_1 e^{-\beta_2 \tau} \|y''(\tau_l) - y_l''\| \quad \forall \tau \in [\tau_l, \tau_{l+1}],$$

which, by (10), leads to

$$\|y''(\tau_{l+1}) - y_l''(\tau_{l+1})\| \leq \beta_1 e^{-\beta_2 \sigma} \|y''(\tau_l) - y_l''\| \leq \frac{1}{2} \|y''(\tau_l) - y_l''\|$$

Consequently, one can use (12) to obtain

$$\begin{aligned} \|y'(\tau_{l+1}) - y_l''(\tau_{l+1})\| &\leq \|y'(\tau_{l+1}) - y_l''(\tau_{l+1})\| + \|y_l''(\tau_{l+1}) - y''(\tau_{l+1})\| \leq \hat{L}_0 \|z' - z''\| + \frac{1}{2} \|y''(\tau_l) - y_l''\| \\ &\leq \hat{L}_0 \|z' - z''\| + \frac{1}{2} [\|y''(\tau_l) - y'(\tau_l)\| + \|y'(\tau_l) - y_l''\|] \leq \frac{3}{2} \hat{L}_0 \|z' - z''\| + \frac{1}{2} \|y'(\tau_l) - y''(\tau_l)\|, \quad l = 0, 1, \dots \end{aligned}$$

As can be readily verified, the estimates above imply (see also (6)) that, for any $l = 0, 1, \dots$,

$$\|y'(\tau_l) - y''(\tau_l)\| \leq 3\hat{L}_0 \|z' - z''\| + \frac{1}{2^l} \|y'_0 - y''_0\| \leq 3\hat{L}_0 \|z' - z''\| + L_0 \|z' - z''\| \leq (3\hat{L}_0 + L_0) \|z' - z''\|. \quad (13)$$

By definition, for any $\tau \in [\tau_l, \tau_{l+1}]$,

$$y'(\tau) = y'(\tau_l) + \int_{\tau_l}^{\tau} f(u'(s), y'(s)) ds, \quad y''(\tau) = y''(\tau_l) + \int_{\tau_l}^{\tau} f(u'(s), y''(s)) ds.$$

Hence, by (13),

$$\|y'(\tau) - y''(\tau)\| \leq (3\hat{L}_0 + L_0) \|z' - z''\| + L_f \int_{\tau_l}^{\tau} \|y'(s) - y''(s)\| ds \quad \forall \tau \in [\tau_l, \tau_{l+1}],$$

where L_f is a Lipschitz constant of $f(u, y)$ (in y). Applying Gronwall- Bellman lemma, one obtains

$$\max_{\tau \in [\tau_l, \tau_{l+1}]} \|y'(\tau) - y''(\tau)\| \leq \hat{L} \|z' - z''\|, \quad l = 0, 1, \dots,$$

where \hat{L} is an appropriately chosen constant. This proves (7) \square

Note that (1) is a stability (dissipativity) type condition. A similar condition was introduced in [15] and was used extensively in dealing with averaging of singularly perturbed control systems with decomposed fast and slow dynamics (see, e.g., [24], [28], [30] and references therein). Emphasize also that (1) is assumed to be satisfied for y', y'' in the invariant set Y_z ($\forall z \in Z$) and it is not assumed to be true for all y', y'' in Y .

Consider a special case in which $F(y) = V^T y$, with $\text{rank}(V) = k$. That is,

$$Y_z = \{y \mid V^T y = z\}. \quad (14)$$

Let us verify the inequality (13) in this case. Let $y' \in Y_{z'}$. That is, $V^T y' = z'$. Let V_c be a $m \times (m - k)$ matrix such that the $m \times m$ matrix $\hat{V} \stackrel{\text{def}}{=} (V, V_c)$ has rank m . Define y'' as the solution of the system

$$V^T y'' = z'', \quad V_c^T y'' = V_c^T y'.$$

Then $y'' \in Y_{z''}$ and

$$\|y' - y''\| = \|(\hat{V}^T)^{-1} ((z' - z'')^T, 0^T)^T\| \leq \|(\hat{V}^T)^{-1}\| \|z' - z''\|.$$

This proves (13) with $L_0 = \|(\hat{V}^T)^{-1}\|$. Note that the set Y_z defined by the right-hand-side of (14) is not compact. However, it can be shown that the solutions of the reduced system are contained in a compact set, say \check{Y} , for any $z \in Z$ (to show this, one may use an argument based on the validity of (2) that is similar to one used in the proof of Theorem 3.1(ii) in [22]). In this case, for the results of Section 2 to be valid, Y should be chosen large enough to contain \check{Y} in its interior and Y_z should be re-defined as the intersection of Y and the set defined by the right-hand-side of (14).

The condition (1) is satisfied if there exist a positive definite matrix Q and a positive number α such that the Jacobian matrix $f'_y(u, y)$ satisfies the inequality

$$v^T [f'_y(u, y)^T Q + Q f'_y(u, y)] v \leq -\alpha \|v\|^2 \quad \forall u \in U, \quad \forall y \in Y_z, \quad \forall v \in \{v \mid V^T v = 0\} \quad \text{and} \quad \forall z \in Z. \quad (15)$$

To further illustrate the latter condition, let us consider an even more special case when

$$f(u, y) = A(u)y, \quad (16)$$

where $A(u)$ is a square matrix function of u and the columns of V are eigenvectors of the matrix $A^T(u)$ corresponding to its eigenvalue 0 that has the multiplicity k . These eigenvectors are assumed to be independent of u as, for example, in the case when, for any $u \in U$, the matrix $A^T(u)$ is equal to the difference of a block-diagonal stochastic matrix and the identity matrix (see Remark 2 below).

Let V_c be as above (that is, it is such that the rank of the matrix $\hat{V} = (V, V_c)$ is equal to m) and let the $m \times (m - k)$ matrix W_c be defined by the equation

$$V^T W_c = 0, \quad V_c^T W_c = I_c \quad \Leftrightarrow \quad W_c = (\hat{V}^T)^{-1} (0, I_c)^T,$$

where I_c is the $(m - k) \times (m - k)$ identity matrix. It can be shown (although we do not do it in the paper) that the inequality (15) is satisfied if there exist a $(m - k) \times (m - k)$ positive definite matrix Q_c and a positive number α_c such that

$$\zeta^T [A_c^T(u) Q_c + Q_c A_c(u)] \zeta \leq -\alpha_c \|\zeta\|^2 \quad \forall u \in U \quad \text{and} \quad \forall \zeta \in \mathbb{R}^{m-k}, \quad (17)$$

where $A_c(u) = V_c^T A(u) W_c$. Note that Q_c and α_c satisfying (17) exist if $A_c(u)$ is independent of u and all its eigenvalues have negative real parts or if $A_c(u)$ is symmetric and its eigenvalues are less than some negative constant for all $u \in U$.

Remark 2 Systems of the form

$$\frac{dy_\epsilon(\tau)}{d\tau} = [A(u(\tau)) + \epsilon B(u(\tau))] y_\epsilon(\tau), \quad (18)$$

(this being a special case of (1), with $f(u, y)$ as in (16) and $g(u, y) = B(u)y$) have been studied extensively in the context of singularly perturbed controlled Markov chains (see, e.g., [52] and references therein). In these studies, $A^T(u)$ was assumed to be equal to the difference of a block-diagonal stochastic matrix and the identity matrix. Each block of $A^T(u)$ was assumed to be irreducible thus ensuring that the zero eigenvalue of $A^T(u)$ has the multiplicity k (provided that the number of blocks is k) and that, corresponding to this eigenvalue, there are k orthogonal eigenvectors that consist of 1s and 0s. In this case, the state variables are the probabilities of the states of the Markov chain and the observables represent the sums of the probabilities of the states that enter the blocks. Note that, with an appropriate change of variables, the system (18) can be decomposed into fast and slow subsystems. However, such change of variables would make the nature of the problem much less transparent and it can also complicate the process of finding a numerical solution; see, e.g., [16], [42] and [52].

The validity of Assumption 2 can also be verified under a controllability condition as stated in the proposition below.

Proposition 4 *Let for any $z \in Z$ and any $y^1, y^2 \in Y_z$, there exist a control $u(\cdot)$ that steers the system (2) from y^1 to y^2 in finite time $T(y^1, y^2, z) \leq T_0$ (T_0 being some positive constant). Then Assumption 2 is satisfied.*

Proof. Let $u^1(\cdot)$ be an arbitrary control and let $y^1(\cdot)$ be a solution of (2) obtained with this control and an initial condition $y^1(0) = y_0^1 \in Y_z$. Also, let $y_0^2 \in Y_z$ and let $u_{1,2}(\tau)$ be the control that steers the system (2) from the point y_0^1 to the point y_0^2 in time $T(y_0^1, y_0^2, z)$. Define the control $u^2(\tau)$ as being equal to $u_{1,2}(\tau)$ for $\tau \in [0, T(y_0^1, y_0^2, z)]$ and as being equal to $u^1(\tau - T(y_0^1, y_0^2, z))$ for $\tau > T(y_0^1, y_0^2, z)$. Denote by $y^2(\cdot)$ the solution of (2) obtained with the control $u^2(\tau)$ and with the initial condition $y^2(0) = y_0^2$. Note that, by construction, $y^2(\tau) = y^1(\tau - T(y_0^1, y_0^2, z)) \forall \tau > T(y_0^1, y_0^2, z)$. Hence, for any continuous $q(u, y)$, the following relationships are valid

$$\begin{aligned} & \left| \frac{1}{S} \int_0^S q(u^1(\tau), y^1(\tau)) d\tau - \frac{1}{S} \int_0^S q(u^2(\tau), y^2(\tau)) d\tau \right| \leq \frac{1}{S} \int_{S-T(y_0^1, y_0^2, z)}^S |q(u^1(\tau), y^1(\tau))| d\tau \\ & + \frac{1}{S} \int_0^{T(y_0^1, y_0^2, z)} |q(u^2(\tau), y^2(\tau))| d\tau + \left| \frac{1}{S} \int_0^{S-T(y_0^1, y_0^2, z)} q(u^1(\tau), y^1(\tau)) d\tau - \frac{1}{S} \int_{T(y_0^1, y_0^2, z)}^S q(u^2(\tau), y^2(\tau)) d\tau \right| \\ & \leq \frac{2M_q}{S}, \quad \text{where} \quad M_q \stackrel{\text{def}}{=} \max_{(u,y) \in U \times Y} |q(u, y)|. \end{aligned}$$

The latter prove (20) with $\beta_q(S) = \frac{2M_q}{S}$, and, thus, the validity of Assumption 2 follows from Proposition 2.

□

Let us give another sufficient condition for Assumption 1 to be satisfied provided that Assumption 2 is valid

Proposition 5 *Let Assumption 2 be valid. Then Assumption 1 will be satisfied if there exists a positive definite matrix Q such that*

$$(f(u, y') - f(u, y''))^T Q (y' - y'') \leq 0 \quad \forall u \in U, \quad \forall y', y'' \in Y. \quad (19)$$

Proof. The proof of the validity of Assumption 1 follows exactly the same lines as in the proof of Proposition 3, with the only difference being the way the estimate (7) is established, where $y'(\tau), y''(\tau)$ are defined as in the aforementioned proof. To prove this estimate, note that, by (19),

$$\frac{d}{d\tau}(y'(\tau) - y''(\tau))^T Q(y'(\tau) - y''(\tau)) = 2(f(u'(\tau), y'(\tau)) - f(u'(\tau), y''(\tau)))^T Q(y'(\tau) - y''(\tau)) \leq 0.$$

Hence,

$$(y'(\tau) - y''(\tau))^T Q(y'(\tau) - y''(\tau)) \leq (y'_0 - y''_0)^T Q(y'_0 - y''_0) \quad \forall \tau \geq 0,$$

which leads to

$$\lambda_{\min} \|y'(\tau) - y''(\tau)\|^2 \leq \lambda_{\max} \|y'_0 - y''_0\|^2 \leq \lambda_{\max} (L_0 \|z' - z''\|)^2 \quad (20)$$

where the latter inequality follows from (6) (λ_{\min} and λ_{\max} being the minimal and the maximum eigenvalues of Q). The estimates (20) imply (7) with $\hat{L} = L_0 \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$. This completes the proof of the proposition. \square

Note that (19) is the “non-expansivity” condition similar to one introduced in [44], and note that it is assumed to be satisfied for all vectors $y', y'' \in Y$ (and not only for ones belonging to the same invariant set). Let us conclude this section with an example.

Example 1. Assume that the perturbed system is of the form

$$\frac{dy_1(\tau)}{d\tau} = u(\tau)y_2(\tau) + \epsilon g_1(u(\tau), y(\tau)), \quad \frac{dy_2(\tau)}{d\tau} = -u(\tau)y_1(\tau) + \epsilon g_2(u(\tau), y(\tau)), \quad (21)$$

where $y(\tau) = (y_1(\tau), y_2(\tau)) \in \mathbb{R}^2$, $u(\tau) \in [-1, 1] \in \mathbb{R}^1$ and $g_i(u, y) : [-1, 1] \times \mathbb{R}^2 \rightarrow \mathbb{R}^1$, $i = 1, 2$ are continuous in (u, y) and Lipschitz continuous in y (uniformly in u). It can be readily verified that the reduced system (that is, the system obtained from (21) with $\epsilon = 0$) satisfies (3) with $F(y) = y_1^2 + y_2^2$. Hence,

$$Y_z = \{(y_1, y_2) \mid y_1^2 + y_2^2 = z\} \quad \forall z \geq 0. \quad (22)$$

Let $Z = [a, b]$, where $0 < a < b$ are some constants. It is easy to check that (13) is valid in this case. Indeed, for any $z', z'' \in Z$ and any $y' \in Y_{z'}$ (note that $\|y'\| = \sqrt{z'}$), take $y'' = \frac{y' \sqrt{z''}}{\sqrt{z'}}$. Then $\|y''\| = \sqrt{z''}$ and $y'' \in Y_{z''}$. Also,

$$\|y' - y''\| = \left| \sqrt{z'} - \sqrt{z''} \right| \leq \frac{|z' - z''|}{2 \min\{\sqrt{z'}, \sqrt{z''}\}} \leq \frac{|z' - z''|}{2\sqrt{a}}.$$

This proves (13) with $L_0 = \frac{1}{2\sqrt{a}}$. It can also be readily seen that, with the use of the control $u(\tau) = 1$, a solution of the reduced system can reach any point in the set Y_z from any other point of this set within a time interval that is less or equal than 2π . Thus, the controllability condition of Proposition 4 is satisfied and, hence, Assumption 2 is valid. One can also verify that the inequality (2) is satisfied with $Q = I$. Indeed,

$$(f(u, y') - f(u, y''))^T (y' - y'') = u [(y'_2 - y''_2), -(y'_1 - y''_1)] [(y'_1 - y''_1), (y'_2 - y''_2)]^T = 0 \quad \forall y', y''.$$

Thus, by Proposition 5, Assumption 1 is valid as well.

4 A case of uncontrolled reduced system

Assume that $f(u, y) = f(y)$. That is, the reduced system (2) has the form

$$\frac{dy(\tau)}{d\tau} = f(y(\tau)) \quad (1)$$

and is uncontrolled. Let us introduce the following assumption about the system (1).

Assumption 3 For any $z \in Z$, the system (1) has a T_z -periodic solution $y_z(\tau) \in Y_z$ (with T_z being uniformly bounded for $z \in Z$) and any other solution of this system with the initial condition in Y_z is a time shift of $y_z(\tau)$.

If Assumption 3 is satisfied, then, for any continuous function $q(y)$ and for any solution $y_{y_0}(\tau)$ of the system (1) satisfying the initial condition $y(0) = y_0 \in Y_z$,

$$\lim_{S \rightarrow \infty} \int_0^S q(y_{y_0}(\tau)) d\tau = \frac{1}{T_z} \int_0^{T_z} q(y_z(\tau)) d\tau, \quad (2)$$

with the convergence being uniform with respect to $y_0 \in Y_z$ and with respect to $z \in Z$. It implies that

$$\lim_{S \rightarrow \infty} \rho(\nu^{y_{y_0}(\cdot), S}, \nu_z) = 0, \quad (3)$$

where $\nu^{y_{y_0}(\cdot), S} \in \mathcal{P}(Y_z)$ is the occupational measure generated by $y_{y_0}(\tau)$ on the interval $[0, S]$ and $\nu_z \in \mathcal{P}(Y_z)$ is the occupational measure generated by the periodic solution $y_z(\tau)$, the latter being defined by the equation

$$\int_{Y_z} q(y) \nu_z(dy) = \frac{1}{T_z} \int_0^{T_z} q(y_z(\tau)) d\tau \quad \forall q(\cdot) \in C \quad (4)$$

Note that ν_z is the invariant measure of the system (1) that is supported on Y_z (see, e.g., [7]). Note also that from the fact that the convergence in (2) is uniform it follows that the convergence in (3) is uniform. That is,

$$\rho(\nu^{y_{y_0}(\cdot), S}, \nu_z) \leq \beta(S) \quad \forall y_0 \in Y_z, \forall z \in Z, \quad \lim_{S \rightarrow \infty} \beta(S) = 0. \quad (5)$$

Proposition 6 Let the reduced system has the form (1) and let Assumption 3 be satisfied. Then:

(i) The set $W(z)$ is presentable in the form

$$W(z) = \{\mu \in \mathcal{P}(U \times Y) : \mu(du, dy) = \sigma(du|y) \nu_z(dy), \quad \sigma(du|\cdot) : Y \rightarrow \mathcal{P}(U)\}, \quad (6)$$

where ν_z is defined by (4) and $\sigma(du|\cdot) : Y \rightarrow \mathcal{P}(U)$ are Borel measurable.

(ii) Assumption 2 is satisfied.

(iii) If, in addition,

$$\frac{|T_{z'} - T_{z''}|}{\max\{T_{z'}, T_{z''}\}} \leq \tilde{L} \|z' - z''\| \quad \forall z', z'' \in Z \quad \text{for some } \tilde{L} = \text{const}, \quad (7)$$

then Assumption 1 is satisfied as well.

Proof of the statements (i) and (ii). By Theorem 2.1(iii) in [23], from (5) it follows that the set

$$\mathcal{V}(z) \stackrel{\text{def}}{=} \{\nu \in \mathcal{P}(Y) : \int_{Y_z} \nabla \phi(y)^T f(y) \nu(dy) = 0 \quad \forall \phi(\cdot) \in C^1, \quad \text{supp}(\nu) \subset Y_z\} \quad (8)$$

is a singleton and that

$$\mathcal{V}(z) = \{\nu_z\}. \quad (9)$$

Note that an arbitrary measure $\mu \in \mathcal{P}(U \times Y)$ can be disintegrated as follows

$$\mu(du, dy) = \sigma(du|y) \nu(dy), \quad (10)$$

where $\nu(dy) = \mu(U, dy)$ and $\sigma(du|y) \in \mathcal{P}(U)$ is such that the integral $\int_U q(u, y) \sigma(du|y)$ is a Borel measurable function on Y for any continuous $q(u, y)$, and

$$\int_{U \times Y} q(u, y) \mu(du, dy) = \int_{y \in Y} \left(\int_U q(u, y) \sigma(du|y) \right) \nu(dy). \quad (11)$$

Note that the Borel measurability of $\int_U q(u, y) \sigma(du|y)$ for any continuous $q(u, y)$ implies the Borel measurability of $\sigma(du|\cdot) : Y \rightarrow \mathcal{P}(U)$.

Proof of (i). Assume that μ belongs to the right-hand-side of (6). That is, $\mu(du, dy) = \sigma(du|y) \nu_z(dy)$. Then $\mu \in W(z)$ (due to (8) and (9)). Conversely, if $\mu \in W(z)$, then using the disintegration (10), one can verify that $\nu \in \mathcal{V}(z)$. Since the latter consists only of one element which is ν_z , this proves that μ belongs to the right hand side of (6). Thus, the validity of (6) is established and (i) is proved.

Proof of (ii). The statement (ii) follows from Proposition 4 since any point in the set Y_z can be reached by a solution of (1) from any other point of this set within a time interval that is less or equal than T_z . \square

The proof of the statement (iii) of Proposition 6 is based on the following lemma.

Lemma 1 *Let, the integral $\int_{Y_z} q(y) \nu_z(dy)$ be continuous in z for any continuous $q(y)$ and let this integral be Lipschitz continuous in z with a Lipschitz constant that can be chosen the same for all Lipschitz continuous $q(y)$ that have a common Lipschitz constant. Then Assumption 1 is satisfied.*

Proof. Assume that the integral $\frac{1}{T_z} \int_0^{T_z} q(y_z(\tau)) d\tau$ is continuous for any continuous $q(\cdot)$. Due to the fact that the map $W(\cdot)$ is convex and compact valued, to prove the continuity of $W(\cdot)$, it is sufficient to show that, if $z_i \rightarrow z$ ($z_i \in Z$), then, for any continuous $q(u, y)$,

$$\lim_{z_i \rightarrow z} \min_{\mu \in W(z_i)} \int_{U \times Y_{z_i}} q(u, y) \mu(du, dy) = \min_{\mu \in W(z)} \int_{U \times Y_z} q(u, y) \mu(du, dy). \quad (12)$$

By (6),

$$\begin{aligned} \min_{\mu \in W(z_i)} \int_{U \times Y_{z_i}} q(u, y) \mu(du, dy) &= \int_{Y_{z_i}} q^*(y) \nu_{z_i}(dy) \\ \min_{\mu \in W(z)} \int_{U \times Y_z} q(u, y) \mu(du, dy) &= \int_{Y_z} q^*(y) \nu_z(dy) \end{aligned}$$

where $q^*(y) \stackrel{\text{def}}{=} \min_{u \in U} q(u, y)$. The function $q^*(y)$ is continuous. Hence,

$$\lim_{z_i \rightarrow z} \int_{Y_{z_i}} q^*(y) \nu_{z_i}(dy) = \int_{Y_z} q^*(y) \nu_z(dy).$$

The latter proves (12) and, thus, the continuity of $W(z)$ is established.

Let L_1 be a Lipschitz constant of the vector function $X(u, y) \stackrel{\text{def}}{=} (h(u, y), r(u, y))$ in y . That is,

$$\|X(u, y') - X(u, y'')\| \leq L_1 \|y' - y''\|. \quad (13)$$

Assume that the integral $\frac{1}{T_z} \int_0^{T_z} q(y_z(\tau)) d\tau$ is Lipschitz continuous with a constant L for any Lipschitz continuous $q(\cdot)$ that have a Lipschitz constant L_1 .

The map $V(z) : Z \rightsquigarrow \mathbb{R}^{k+1}$ defined in (14) is convex and compact valued. Hence to prove (15), it is sufficient to prove that, for any $v \in \mathbb{R}^{k+1}$ such that $\|v\| = 1$,

$$\left| \min_{\mu \in W(z')} \int_{U \times Y_{z'}} v^T X(u, y) \mu(du, dy) - \min_{\mu \in W(z'')} \int_{U \times Y_{z''}} v^T X(u, y) \mu(du, dy) \right| \leq L \|z' - z''\| \quad (14)$$

(see, e.g., Lemma P.2.9, p. 207 in [21]). By (6),

$$\begin{aligned} \min_{\mu \in W(z')} \int_{U \times Y_{z'}} v^T X(u, y) \mu(du, dy) &= \int_{Y_{z'}} q_v^*(y) \nu_{z'}(dy), \\ \min_{\mu \in W(z'')} \int_{U \times Y_{z''}} v^T X(u, y) \mu(du, dy) &= \int_{Y_{z''}} q_v^*(y) \nu_{z''}(dy), \end{aligned}$$

where $q_v^*(y) \stackrel{\text{def}}{=} \min_{u \in U} \{v^T X(u, y)\}$. Note that from (13) it follows that $|v^T X(u, y') - v^T X(u, y'')| \leq L_1 \|y' - y''\|$, which, in turn, implies $|q_v^*(y') - q_v^*(y'')| \leq L_1 \|y' - y''\|$. Thus

$$\left| \int_{Y_{z'}} q_v^*(y) \nu_{z'}(dy) - \int_{Y_{z''}} q_v^*(y) \nu_{z''}(dy) \right| \leq L \|z' - z''\|.$$

This proves (14), and, thus, the proof of the lemma is completed. \square

Proof of the statement (iii) of Proposition 6. To prove the statement, it is sufficient to verify that the conditions of Lemma 1 are satisfied.

Let $z', z'' \in Z$. Let $y_{z'}(\tau)$ be a solution of (1) such that $y_{z'}(0) \in Y_{z'}$ and let $y_{z''}(\tau)$ be the solution of (1) that satisfies the initial condition $y_{z''}(0) = \operatorname{argmin}_{y \in Y_{z''}} \{\|y - y_{z'}(0)\|\}$. From (13) it follows that $\|y_{z''}(0) - y_{z'}(0)\| \leq L_0 \|z'' - z'\|$. Using this estimate and applying Gronwall-Bellman Lemma, one can deduce that

$$\|y_{z''}(\tau) - y_{z'}(\tau)\| \leq \hat{L}_0 \|z'' - z'\| \quad \forall \tau \in [0, \max\{T_{z'}, T_{z''}\}], \quad (15)$$

where \hat{L}_0 is an appropriately chosen constant.

Assume (without loss of generality) that $T_{z''} \geq T_{z'}$. For an arbitrary continuous function $q(y)$, one can write down the following inequalities

$$\begin{aligned} & \left| \frac{1}{T_{z'}} \int_0^{T_{z'}} q(y_{z'}(\tau)) d\tau - \frac{1}{T_{z''}} \int_0^{T_{z''}} q(y_{z''}(\tau)) d\tau \right| \leq \frac{1}{T_{z'}} \int_0^{T_{z'}} |q(y_{z'}(\tau)) - q(y_{z''}(\tau))| d\tau \\ & + \left| \frac{1}{T_{z''}} - \frac{1}{T_{z'}} \right| \int_0^{T_{z'}} |q(y_{z''}(\tau))| d\tau + \frac{1}{T_{z''}} \int_{T_{z'}}^{T_{z''}} |q(y_{z''}(\tau))| d\tau \leq \kappa_q(\hat{L}_0 \|z'' - z'\|) + 2M_q \frac{|T_{z'} - T_{z''}|}{T_{z''}} \\ & \leq \kappa_q(\hat{L}_0 \|z'' - z'\|) + 2M_q \tilde{L} \|z' - z''\| \end{aligned}$$

where (15) and, subsequently, (7) have been taken into account and where $\kappa_q(\cdot)$ is the modulus of continuity of $q(\cdot)$ on Y ($\lim_{\theta \rightarrow 0} \kappa_q(\theta) = 0$) and $M_q = \max_{y \in Y} |q(y)|$. By (4), the latter imply

$$\left| \int_{Y_{z'}} q(y) \nu_{z'}(dy) - \int_{Y_{z''}} q(y) \nu_{z''}(dy) \right| \leq \kappa_q(\hat{L}_0 \|z'' - z'\|) + 2M_q \tilde{L} \|z' - z''\|, \quad (16)$$

which proves that the integral $\int_{Y_z} q(y) \nu_z(dy)$ is continuous in z . Also, it proves that this integral is Lipschitz continuous with the constant $L_1 \hat{L}_0 + 2M_q \tilde{L}$ for any Lipschitz continuous $q(\cdot)$ that has a Lipschitz constant L_1 (since $\kappa_q(\hat{L}_0 \|z'' - z'\|) \leq L_1 \hat{L}_0 \|z'' - z'\|$ in this case). This completes the proof of the statement. \square

Let us consider two examples in which the conditions of Proposition 6 are satisfied.

Example 1 (continued). Consider system (21) with $u(\tau) \equiv 1$ (an uncontrolled version of the system). As one can see, this system satisfies Assumption 3 with $T_z = 2\pi$ for all $z > 0$. We have already checked that (13) is satisfied on $Z = [a, b]$ ($0 < a < b$). Also, one can see that (7) is valid with $\tilde{L} = 0$.

Example 2 (perturbed Lotka-Volterra equations). Consider the system

$$\frac{dy_1(\tau)}{d\tau} = -y_1(\tau) + y_1(\tau)y_2(\tau) - \epsilon u(\tau)y_1(\tau), \quad \frac{dy_2(\tau)}{d\tau} = y_2(\tau) - y_1(\tau)y_2(\tau), \quad (17)$$

where $y = (y_1, y_2) \in \mathbb{R}^2$ and controls $u(\cdot)$ are scalar functions of time satisfying the inclusion $u(\tau) \in [0, 1]$. A system similar to (17) but with no controls and with the presence of small random perturbations was considered in [34]. The reduced system (1) is the Lotka-Volterra system

$$\frac{dy_1(\tau)}{d\tau} = -y_1(\tau) + y_1(\tau)y_2(\tau), \quad \frac{dy_2(\tau)}{d\tau} = y_2(\tau) - y_1(\tau)y_2(\tau), \quad (18)$$

and it is known that (3) is satisfied with

$$F(y) = \ln y_2(t) - y_2(t) + \ln y_1(t) - y_1(t). \quad (19)$$

Note that it can be shown that the system (18) does not satisfy the non-expansivity condition (19).

For any $z < -2$, the set $Y_z \stackrel{\text{def}}{=} \{y \in \mathbb{R}^2 : F(y) = z\}$ is a closed curve, the solutions of the reduced system being periodic and moving along this curve. Two such curves (for $z = -3$ and $z = -2.05$) are depicted in Fig. 1 in Section 6 below. For $z = -2$, the curve is degenerated into the equilibrium point $(1, 1)$. Note that $T_{z''} < T_{z'}$ if $z'' > z'$, where T_z is the period of the solutions of (18) that are contained in Y_z (see, e.g., Theorem 2 in [49]).

Let

$$Z = [-3, -2.05], \quad Y = [a_1, b_1] \times [a_2, b_2], \quad (20)$$

where $0 < a_i < b_i$, $i = 1, 2$, are such that $Y_z \subset Y \forall z \in [-3, -2.05]$. Assumption 3 is obviously satisfied in this case, and it can be readily verified that (13) is valid (this follows from the fact that $\min_{y \in Y_z} \|F'(y)\| \geq \alpha > 0 \forall z \in Z$). Since T_z is a monotone decreasing, $T_z \geq T_{-2.05} > 0 \forall z \in Z$. Hence, to verify (7), it is sufficient to establish that T_z is Lipschitz continuous on Z . The proof of this is, however, quite technical, and we do not present it in the paper.

In Section 6, we will return to this example in the context of numerical construction of asymptotically near optimal solution of an optimal control problem considered on the trajectories of the perturbed system (17). In the remainder of this section we will establish statements similar to (i) and (ii) of Proposition 6 under a weaker assumption.

Assumption 4 *There exists a solution $\bar{y}(t) \in Y_z$ of the reduced system (1) such that the following two conditions are satisfied:*

(i) *The solution $\bar{y}(\cdot)$ generates an occupational measure ν_z . That is, for any continuous function $q(y)$,*

$$\left| \frac{1}{S} \int_0^S q(\bar{y}(t)) dt - \int_{Y_z} q(y) \nu_z(dy) \right| \leq \kappa_q(S), \quad \text{where} \quad \lim_{S \rightarrow \infty} \kappa_q(S) = 0. \quad (21)$$

(ii) *There exist positive functions $\alpha_1(S)$ and $\alpha_2(S)$,*

$$\lim_{S \rightarrow \infty} \alpha_i(S) = 0, \quad i = 1, 2, \quad (22)$$

such that any other solution $y(t) \in Y_z$ of the reduced system (2) satisfies the inequality

$$\max_{\tau \in [0, S]} \|y(\tau + \theta_S) - \bar{y}(\tau)\| \leq \alpha_1(S), \quad (23)$$

for some θ_S such that

$$\frac{\theta_S}{S} \in [0, \alpha_2(S)]. \quad (24)$$

Note that Assumption 4 is satisfied if Assumption 3 is satisfied, in which case one may take $\alpha_1(S) = 0$ and $\alpha_2(S) = \frac{T_z}{S}$, where T_z is the period of the solutions staying in Y_z .

Proposition 7 *Let Assumption 4 be satisfied. Then the statements (i) and (ii) of Proposition 6 are valid.*

Proof. From Assumption 4(ii) it follows that, for any Lipschitz continuous function $q(y)$ and any solution $y(t)$ of the system (1),

$$\begin{aligned} \left| \frac{1}{S} \int_0^S q(y(\tau)) d\tau - \frac{1}{S} \int_0^S q(\bar{y}(\tau)) d\tau \right| &\leq \frac{1}{S} \int_0^{S-\theta_S} |q(y(\tau + \theta_S)) - q(\bar{y}(\tau))| d\tau + \frac{2M_q \theta_S}{S} \\ &\leq L_q \alpha_1(S) + \frac{2M_q \theta_S}{S} \leq L_q \alpha_1(S) + 2M_q \alpha_2(S) \stackrel{\text{def}}{=} \alpha_q(S), \end{aligned} \quad (25)$$

where L_q is a Lipschitz constant of $q(\cdot)$ and $M_q \stackrel{\text{def}}{=} \max_{y \in Y} |q(y)|$. Then

$$\left| \frac{1}{S} \int_0^S q(y(\tau)) d\tau - \int_{Y_z} q(y) \nu_z(dy) \right| \leq \alpha_q(S) + \kappa_q(S),$$

where $\kappa_q(S)$ is as in (21). The latter implies that the occupational measure generated by any solution staying in Y_z converges to ν_z . That is, (5) is valid. From this point, the proofs of (i) follows exactly the same lines as those of Proposition 6(i).

To prove (ii), let us show that given an arbitrary pair $(u^1(\cdot), y^1(\cdot))$, where $y^1(\cdot)$ is a solution of (1) with $y^1(0) \in Y_z$ and given another solution $y^2(\cdot)$ of (1) with $y^2(0) \in Y_z$, there exists a control $u^2(\cdot)$ such that, for any Lipschitz continuous function $q(u, y)$, the estimate (20) is valid (see Proposition 2). To this end, define, first, $\bar{u}(\tau)$ to be equal to $u^1(\tau + \theta_S)$ for $\tau \in [0, S - \theta_S]$ and to be equal to an arbitrary $u \in U$ on the interval $(S - \theta_S, S]$. Then one can obtain

$$\begin{aligned} & \left| \frac{1}{S} \int_0^S q(u^1(\tau), y^1(\tau)) d\tau - \frac{1}{S} \int_0^S q(\bar{u}(\tau), \bar{y}(\tau)) d\tau \right| \\ & \leq \frac{1}{S} \int_0^{S-\theta_S} |q(u^1(\tau + \theta_S), y^1(\tau + \theta_S)) - q(u^1(\tau + \theta_S), \bar{y}(\tau))| d\tau + \frac{2M_q\theta_S}{S} \\ & \leq L_q \alpha_1(S) + \frac{2M_q\theta_S}{S} \leq L_q \alpha_1(S) + 2M_q \alpha_2(S) \stackrel{\text{def}}{=} \alpha_q(S), \end{aligned} \quad (26)$$

where L_q is a Lipschitz constant of $q(\cdot)$ and $M_q \stackrel{\text{def}}{=} \max_{(u,y) \in U \times Y} |q(u, y)|$. Define now $u^2(\tau)$ to be equal to an arbitrary $u \in U$ for $\tau \in [0, \theta_S]$ and to be equal to $\bar{u}(\tau - \theta_S)$ for $\tau \in [\theta_S, S]$. Then, similarly to (25), one obtains

$$\begin{aligned} & \left| \frac{1}{S} \int_0^S q(u^2(\tau), y^2(\tau)) d\tau - \frac{1}{S} \int_0^S q(\bar{u}(\tau), \bar{y}(\tau)) d\tau \right| \\ & \leq \frac{1}{S} \int_0^{S-\theta_S} |q(u^2(\tau + \theta_S), y^2(\tau + \theta_S)) - q(\bar{u}(\tau), \bar{y}(\tau))| d\tau + \frac{2M_q\theta_S}{S} \\ & = \frac{1}{S} \int_0^{S-\theta_S} |q(\bar{u}(\tau), y^2(\tau + \theta_S)) - q(\bar{u}(\tau), \bar{y}(\tau))| d\tau + \frac{2M_q\theta_S}{S} \\ & \leq L_q \alpha_1(S) + \frac{2M_q\theta_S}{S} \leq L_q \alpha_1(S) + 2M_q \alpha_2(S) = \alpha_q(S), \end{aligned} \quad (27)$$

The validity of (26) and (27) implies (20) with $\beta_q(S) = 2\alpha_q(S)$. \square

5 Optimal and near optimal average control generating families

In this section, we discuss a way of characterization and construction of optimal or near optimal average control generating (ACG) families (a concept introduced in [26] and [30] for a different class of problems) and, in the next section, we will demonstrate how these can be used for the construction of asymptotically optimal or asymptotically near optimal solutions of the perturbed problem.

Let $u_z(\tau)$ be a parameterized by z family of controls and let $y_z(\tau) \in Y_z$ be a family of solutions of the reduced system (2) obtained with these controls such that, for any $z \in Z$, the pair $(u_z(\cdot), y_z(\cdot))$ generates the occupational measure $\mu_z(du, dy)$ (see (14)), with the integral $\int_{U \times Y} q(u, y) \mu_z(du, dy)$ being a measurable function of z , and

$$|S^{-1} \int_0^S q(u_z(\tau), y_z(\tau)) d\tau - \int_{U \times Y} q(u, y) \mu_z(du, dy)| \leq \alpha_q(S) \quad \forall z \in Z, \quad \lim_{S \rightarrow \infty} \alpha_q(S) = 0 \quad (1)$$

for any continuous $q(u, y)$. Note that the estimate (1) is valid if $y_z(\cdot)$ and $u_z(\cdot)$ are T_z -periodic, with T_z being uniformly bounded on Z . Note also that (similarly to the proof of Proposition 1) it can be verified that

$$\mu_z \in W(z) \quad \forall z \in Z. \quad (2)$$

Definition 1 The family $(u_z(\cdot), y_z(\cdot))$ will be called *average control generating (ACG)* if the system

$$z'(t) = \tilde{h}(\mu_{z(t)}), \quad z(0) = z_0, \quad (3)$$

where

$$\tilde{h}(\mu_z) = \int_{U \times Y} h(u, y) \mu_z(du, dy), \quad (4)$$

has a unique solution $z(t) \in Z \forall t \in [0, \infty)$ (thus, the pair $(\mu_{z(t)}, z(t))$ is a solution of the averaged system (6), (7)).

Definition 2 An ACG family $(u_z(\cdot), y_z(\cdot))$ is called *optimal* if the pair $(\mu_{z(t)}, z(t))$, where $z(t)$ is the solution of (3), is an optimal solution of the averaged problem (8). That is,

$$\int_0^{+\infty} e^{-Ct} \tilde{r}(\mu_{z(t)}) dt = \tilde{R}^*(z_0). \quad (5)$$

An ACG family $(u_z(\cdot), y_z(\cdot))$ is called α -near optimal ($\alpha > 0$) if

$$\int_0^{+\infty} e^{-Ct} \tilde{r}(\mu_{z(t)}) dt \leq \tilde{R}^*(z_0) + \alpha. \quad (6)$$

Let Φ stand for the graph of $W(\cdot)$ (see (2)):

$$\Phi \stackrel{\text{def}}{=} \{(\mu, z) : \mu \in W(z), \quad z \in Z\} \subset \mathcal{P}(U \times Y) \times Z. \quad (7)$$

and let the set $\tilde{\mathcal{D}}(z_0)$ be defined by the equation

$$\tilde{\mathcal{D}}(z_0) \stackrel{\text{def}}{=} \{p \in \mathcal{P}(\Phi) : \int_{\Phi} (\nabla \psi(z)^T \tilde{h}(\mu) + C(\psi(z_0) - \psi(z))) p(d\mu, dz) = 0 \quad \forall \psi(\cdot) \in C^1(\mathbb{R}^k)\} \quad (8)$$

In what follows, it is always assumed that

$$\tilde{\mathcal{D}}(z_0) \neq \emptyset. \quad (9)$$

Note that, the latter is satisfied if (24) is valid (since the discounted occupational measure generated by any solution $(\mu(\cdot), z(\cdot))$ of the averaged system is contained in $\tilde{\mathcal{D}}(z_0)$; see Lemma 2.1 in [29]).

Let us consider the following optimization problem.

$$\min_{p \in \tilde{\mathcal{D}}(z_0)} \int_{\Phi} \tilde{r}(\mu) p(d\mu, dz) \stackrel{\text{def}}{=} \tilde{a}^*(z_0). \quad (10)$$

Note that this problem belongs to the class of infinite dimensional linear programming (IDL) problems (since both the objective function and the constraints are linear in the “decision variable” p). The optimal solution of the problem (10) exists (since, as can be readily verified, $\tilde{\mathcal{D}}(z_0)$ is weak* compact).

The IDLP problem (10) is closely related to the averaged optimal control problem (8). In particular, the optimal values of these problem are related by the inequality (see Proposition 2.2 in [28] and Lemma 2.1 in [29])

$$C\tilde{R}^*(z_0) \geq \tilde{a}^*(z_0), \quad (11)$$

which, under certain conditions, becomes the equality

$$C\tilde{R}^*(z_0) = \tilde{a}^*(z_0). \quad (12)$$

In fact, the equality (12) is valid if, by allowing the use of relaxed controls, we would not improve the optimal value of the averaged problem (see Theorem 2.2 and Corollary 2.3 in [29]). It seems likely that the averaged problem does satisfy this property but we do not verify it in the paper.

Consider the “maximin” problem

$$\sup_{\zeta(\cdot) \in C^1(\mathbb{R}^k)} \min_{z \in Z} \{ \tilde{H}(\nabla \zeta(z), z) + C(\zeta(z_0) - \zeta(z)) \} = \tilde{a}^*(z_0), \quad (13)$$

where \sup is sought over all continuously differentiable functions $\zeta(\cdot) : \mathbb{R}^k \rightarrow \mathbb{R}^1$ and $\tilde{H}(p, z)$ stands for the Hamiltonian of the averaged system,

$$\tilde{H}(p, z) \stackrel{\text{def}}{=} \min_{\mu \in W(z)} \{ \tilde{r}(\mu) + p^T \tilde{h}(\mu) \}. \quad (14)$$

The problem (13) is dual with respect to the IDLP problem (10), the duality relationships between the two problems include the equality of their optimal values (see Theorem 3.1 in [28]). For convenience, (13) will be referred to as the *averaged dual problem*.

The right-hand-side in (14) can be rewritten in the form of the problem (see (4) and (9))

$$\min_{\mu \in W(z)} \left\{ \int_{U \times Y} [r(u, y) + p^T h(u, y)] \mu(du, dy) \right\} = \tilde{H}(p, z), \quad (15)$$

which is of the IDLP class as well (since its objective functions and the constraints defining $W(z)$ are linear in μ ; see (2)). The problem dual with respect to (15) is the maximin problem

$$\sup_{\eta(\cdot) \in C^1(\mathbb{R}^m)} \min_{(u, y) \in U \times Y_z} \{ r(u, y) + p^T h(u, y) + \nabla \eta(y)^T f(u, y) \} = \tilde{H}(p, z), \quad (16)$$

where \sup is sought over all continuously differentiable functions $\eta(\cdot) : \mathbb{R}^m \rightarrow \mathbb{R}^1$. The optimal values of the problems (15) and (16) are equal, this being one of the duality relationships between these two problems (see Theorem 4.1 in [18]). The problem (16) will be referred to as the *associated dual problem*.

Assumption 5

(i) A solution of the averaged dual problem exists. That is, there exists $\zeta^*(\cdot) \in C^1(\mathbb{R}^k)$ such that

$$\min_{z \in Z} \{ \tilde{H}(\nabla \zeta^*(z), z) + C(\zeta^*(z_0) - \zeta^*(z)) \} = \tilde{a}^*(z_0). \quad (17)$$

(ii) A solution of the associated dual problem considered with $p = \nabla \zeta^*(z)$ (where $\zeta^*(z)$ is a solution of the averaged dual problem) exists for any $z \in Z$. That is, there exists $\eta_z^*(\cdot) \in C^1(\mathbb{R}^m)$ such that

$$\min_{(u, y) \in U \times Y_z} \{ r(u, y) + \nabla \zeta^*(z)^T h(u, y) + \nabla \eta_z^*(y)^T f(u, y) \} = \tilde{H}(\nabla \zeta^*(z), z) \quad \forall u \in U \quad \forall z \in Z. \quad (18)$$

Note that from that (14) and (17) it follows that

$$\min_{(\mu, z) \in \Phi} \{ \tilde{r}(\mu) + \nabla \zeta^*(z)^T \tilde{h}(\mu) + C(\zeta(z_0) - \zeta(z)) \} = \tilde{a}^*(z_0), \quad (19)$$

and from (17) and (18) it follows that

$$\min_{z \in Z} \min_{(u, y) \in U \times Y_z} \{ r(u, y) + \nabla \zeta^*(z)^T h(u, y) + \nabla \eta_z^*(y)^T f(u, y) + C(\zeta^*(z_0) - \zeta^*(z)) \} = \tilde{a}^*(z_0). \quad (20)$$

The result stated below gives a sufficient condition for an ACG family $(u_z(\cdot), y_z(\cdot))$ to be optimal. This also becomes a necessary condition of optimality provided that the ACG family is periodic, that is,

$$(u_z(\tau + T_z), y_z(\tau + T_z)) = (u_z(\tau), y_z(\tau)) \quad \forall \tau \geq 0 \quad (21)$$

for some $T_z > 0$.

Proposition 8 *Let Assumption 5 be satisfied. Then, for an ACG family $(u_z(\cdot), y_z(\cdot))$ to be optimal and for the equality (12) to be valid it is sufficient and, provided that (21) is satisfied, also necessary that there exist sets $P_t \subset [0, \infty)$, $A \subset [0, \infty)$ such that*

$$\text{meas}\{[0, \infty) \setminus P_t\} = 0 \quad \forall t \in A \quad \text{and} \quad \text{meas}\{[0, \infty) \setminus A\} = 0, \quad (22)$$

($\text{meas}\{\cdot\}$ stands for the Lebesgue measure of the corresponding set) and such that

$$\begin{aligned} & r(u_{z(t)}(\tau), y_{z(t)}(\tau)) + \nabla \zeta^*(z(t))^T h(u_{z(t)}(\tau), y_{z(t)}(\tau)) + \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau)) \\ & + C(\zeta^*(z_0) - \zeta^*(z(t))) = \tilde{a}^*(z_0) \quad \forall \tau \in P_t, \quad \forall t \in A, \end{aligned} \quad (23)$$

where $z(t)$ is the solution of (3) and $\mu_{z(t)}$ in (3) is the occupational measure generated by $(u_{z(t)}(\cdot), y_{z(t)}(\cdot))$.

Proof. We assume first that (23) is satisfied and show that the ACG family $(u_z(\cdot), y_z(\cdot))$ is optimal and that (12) is valid. By integrating (23) over $\tau \in [0, S]$, dividing the result by S and passing to the limit with $S \rightarrow \infty$ (and also having in mind (4), (9) and (1)), one obtains

$$\tilde{r}(\mu_{z(t)}) + \nabla \zeta^*(z(t))^T \tilde{h}(\mu_{z(t)}) + C(\zeta^*(z_0) - \zeta^*(z(t))) = \tilde{a}^*(z_0) \quad \forall t \in A, \quad (24)$$

where it has been taken into account that

$$\lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S \nabla \eta_{z(t)}^*(u_{z(t)}(\tau), y_{z(t)}(\tau))^T f(y_{z(t)}(\tau)) d\tau = \lim_{S \rightarrow \infty} \frac{1}{S} \left(\eta_{z(t)}^*(y_{z(t)}(S)) - \eta_{z(t)}^*(y_{z(t)}(0)) \right) = 0.$$

By multiplying (24) by e^{-Ct} and integrating the result over t from 0 to ∞ , one obtains

$$\int_0^\infty e^{-Ct} \tilde{r}(\mu_{z(t)}) dt = \frac{1}{C} \tilde{a}^*(z_0) \quad (25)$$

where it has been taken into account that

$$\int_0^\infty e^{-Ct} \left(\nabla \zeta^*(z(t))^T \tilde{h}(\mu_{z(t)}) + C(\zeta^*(z_0) - \zeta^*(z(t))) \right) dt = \int_0^\infty \frac{d}{dt} (e^{-Ct} \zeta^*(z(t))) + \zeta^*(z_0) = 0. \quad (26)$$

Due to (11), the validity of (25) implies both the optimality of the ACG under consideration and the validity of (12).

Assume now the ACG family is optimal and (12) is valid. Then (25) is true, which along with (26) implies that

$$\int_0^\infty e^{-Ct} \left(\tilde{r}(\mu_{z(t)}) + \nabla \zeta^*(z(t))^T \tilde{h}(\mu_{z(t)}) + C(\zeta^*(z_0) - \zeta^*(z(t))) - \tilde{a}^*(z_0) \right) dt = 0. \quad (27)$$

Due to (19), from (27) it follows that

$$\tilde{r}(\mu_{z(t)}) + \nabla \zeta^*(z(t))^T \tilde{h}(\mu_{z(t)}) + C(\zeta^*(z_0) - \zeta^*(z(t))) = \tilde{a}^*(z_0) \quad \forall t \in A, \quad (28)$$

where $A \subset [0, \infty)$ is such that $\text{meas}\{[0, \infty) \setminus A\} = 0$. Having in mind the fact that, under the periodicity condition (21),

$$\tilde{r}(\mu_{z(t)}) = \frac{1}{T_{z(t)}} \int_0^{T_{z(t)}} r(u_{z(t)}(\tau), y_{z(t)}(\tau)) d\tau, \quad \tilde{h}(\mu_{z(t)}) = \frac{1}{T_{z(t)}} \int_0^{T_{z(t)}} h(u_{z(t)}(\tau), y_{z(t)}(\tau)) d\tau$$

and

$$\int_0^{T_{z(t)}} \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau)) d\tau = \eta_{z(t)}^*(y_{z(t)}(T_{z(t)})) - \eta_{z(t)}^*(y_{z(t)}(0)) = 0,$$

one can rewrite (28) in the form

$$\begin{aligned} \frac{1}{T_{z(t)}} \int_0^{T_{z(t)}} & (r(u_{z(t)}(\tau), y_{z(t)}(\tau)) + \nabla \zeta^*(z(t))^T h(u_{z(t)}(\tau), y_{z(t)}(\tau)) + \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau)) \\ & + C(\zeta^*(z_0) - \zeta^*(z(t)) - \tilde{a}^*(z_0)) d\tau = 0 \quad \forall t \in A. \end{aligned} \quad (29)$$

The latter along with (20) imply the validity of the equality (23) for almost all $\tau \in [0, T_{z(t)}]$. This, in turn, implies the validity of (23) for $t \in P_t$, where P_t is such that $\text{meas}\{[0, \infty) \setminus P_t\} = 0$. This completes the proof. \square

Corollary 3 *Let Assumption 5 be satisfied and the equality (12) be valid. Then for an ACG family satisfying the periodicity condition (21) to be optimal it is necessary that, for any $\tau \in P_t$ and any $t \in A$,*

$$u_{z(t)}(\tau) \in \text{Argmin}_{u \in U} \{r(u, y_{z(t)}(\tau)) + \nabla \zeta^*(z(t))^T h(u, y_{z(t)}(\tau)) + \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u, y_{z(t)}(\tau))\} \quad (30)$$

Proof. If the ACG $(u_z(\cdot), y_z(\cdot))$ is optimal, then (23) is satisfied. The latter and (20) imply that

$$(u_{z(t)}(\tau), y_{z(t)}(\tau)) \in \text{Argmin}_{(u, y) \in U \times Y} \{r(u, y) + \nabla \zeta^*(z(t))^T h(u, y) + \nabla \eta_{z(t)}^*(y)^T f(u, y)\} \quad \forall \tau \in P_t, \quad \forall t \in A,$$

which, in turn, implies (30). \square

Having in mind (30), one may pose a question whether the feedback control $u^*(y, z)$ defined as a minimizer

$$u^*(y, z) \stackrel{\text{def}}{=} \text{argmin}_{u \in U} \{r(u, y) + \nabla \zeta^*(z)^T h(u, y) + \eta_z^*(y)^T f(u, y)\} \quad (31)$$

generates an optimal ACG family.

Definition 3 *A feedback control function $u^*(y, z)$ is referred to as that generating an optimal (α -near optimal) ACG family if a pair $(u_z^*(\cdot), y_z^*(\cdot))$, where $y_z^*(\tau) \in Y_z$ is a solution of the system*

$$\frac{dy(\tau)}{d\tau} = f(u^*(y(\tau), z), y(\tau)) \quad (32)$$

and $u_z^(\tau) = u^*(y_z^*(\tau), z)$, generates an optimal (α -near optimal) ACG family.*

Proposition 9 *Let Assumption 5 be satisfied and the equality (12) be valid. Let $u^*(y, z)$ be as in (31). Assume that, for any $z \in Z$, the set of occupational measures $W^*(z)$ that are generated by the pairs $(u_z^*(\cdot), y_z^*(\cdot))$, where $y_z^*(\tau)$ is a solution of the system (32) and $u_z^*(\tau) = u^*(y_z^*(\tau), z)$, is a singleton, $W^*(z) = \{\mu_z^*\}$. Assume also that an optimal ACG family $(u_z(\cdot), y_z(\cdot))$ satisfying the periodicity condition (21) exists and the problem*

$$\min_{u \in U} \{r(u, y_{z(t)}(\tau)) + \nabla \zeta^*(z(t))^T h(u, y_{z(t)}(\tau)) + \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u, y_{z(t)}(\tau))\} \quad (33)$$

has a unique minimizer for almost all $\tau \in [0, \infty)$ and almost all $t \in [0, \infty)$, where $z(t)$ is the solution of the system (3). Then $u^(y, z)$ generates an optimal ACG family.*

Proof. By Corollary 3, $u_{z(t)}(\tau)$ is a minimizer in the problem (33) for all $\tau \in P_t$ and for all $t \in A$. Due to the fact that the minimizer of this problem is unique,

$$u_{z(t)}(\tau) = u^*(y_{z(t)}(\tau), z(t))$$

for almost all $\tau \in P_t$ and almost all $t \in A$. Since $\text{meas}\{[0, \infty) \setminus P_t\} = 0$, $y_{z(t)}(\tau)$ is a solution of (32) and, thus, the occupational measure $\mu_{z(t)}$ generated by $(u_z(\cdot), y_z(\cdot))$ is an element of $W^*(z(t))$ for almost all $t \in A$. The set $W^*(z(t))$ consists of only one element $\mu_{z(t)}^*$ and, hence,

$\mu_{z(t)} = \mu_{z(t)}^*$ for almost all $t \in [0, \infty)$ (remind that $meas\{[0, \infty) \setminus A\} = 0\}$). This proves the proposition. \square

Solutions of the averaged and associated dual problems may not exist (that is, Assumption 5 may not be satisfied) and, even if they exist, finding these solutions analytically is hardly possible. A way of finding approximate solutions of the dual problems numerically was proposed in [30] for a different class of problems. The approach of [30] is applicable for the class of problems under consideration as well. Below, we briefly describe a way this approach can be used for construction of feedback controls that generate near optimal ACG families.

Let $\psi_i(\cdot) \in C^1(\mathbb{R}^k)$, $i = 1, 2, \dots$, be a sequence of functions such that any $\zeta(\cdot) \in C^1(\mathbb{R}^k)$ and its gradient are simultaneously approximated by a linear combination of $\psi_i(\cdot)$ and their gradients. Also, let $\phi_i(\cdot) \in C^1(\mathbb{R}^m)$, $i = 1, 2, \dots$, be a sequence of functions such that any $\eta(\cdot) \in C^1(\mathbb{R}^m)$ and its gradient are simultaneously approximated by a linear combination of $\phi_i(\cdot)$ and their gradients. Examples of such sequences are monomials $z_1^{i_1} \dots z_k^{i_k}$, $i_1, \dots, i_k = 0, 1, \dots$ and, respectively, $y_1^{i_1} \dots y_m^{i_m}$, $i_1, \dots, i_m = 0, 1, \dots$, with z_j ($j = 1, \dots, k$) and y_l ($l = 1, \dots, m$) standing for the components of z and y (see, e.g., [38]).

Let us introduce the following notations:

$$W_M(z) \stackrel{\text{def}}{=} \{\mu \in \mathcal{P}(U \times Y) : \int_{U \times Y} \nabla \phi_i(y)^T f(u, y, z) \mu(du, dy) = 0, \quad i = 1, \dots, M\}, \quad (34)$$

$$\Phi_M \stackrel{\text{def}}{=} \{(\mu, z) : \mu \in W_M(z), \quad z \in Z\} \subset \mathcal{P}(U \times Y) \times Z, \quad (35)$$

$$\tilde{H}_M(p, z) \stackrel{\text{def}}{=} \min_{\mu \in W_M(z)} \{\tilde{r}(\mu) + p^T \tilde{h}(\mu)\} \quad (36)$$

(compare with (2), (7) and (14)) and let us consider the maximin problem

$$\sup_{\zeta(\cdot) \in \mathcal{Q}_N} \min_{z \in Z} \{\tilde{H}_M(\nabla \zeta(z), z) + C(\zeta(F(y_0)) - \zeta(z))\} \stackrel{\text{def}}{=} \tilde{a}_{M,N}(z_0), \quad (37)$$

where \sup is over the functions from the finite dimensional space $\mathcal{Q}_N \subset C^1(\mathbb{R}^k)$,

$$\mathcal{Q}_N \stackrel{\text{def}}{=} \{\zeta(\cdot) \in C^1(\mathbb{R}^k) : \zeta(z) = \sum_{i=1}^N \lambda_i \psi_i(z), \quad \lambda = (\lambda_i) \in \mathbb{R}^N\} \quad (38)$$

(compare with (13)). The problem (38) is dual with respect to the semi-infinite LP problem

$$\min_{p \in \tilde{\mathcal{D}}_{M,N}(z_0)} \int_{F_M} \tilde{r}(\mu) p(d\mu, dz) = \tilde{a}_{M,N}(z_0), \quad (39)$$

where

$$\tilde{\mathcal{D}}_{M,N}(z_0) \stackrel{\text{def}}{=} \{p \in \mathcal{P}(\Phi_M) : \int_{\Phi_M} (\nabla \psi_i(z)^T \tilde{h}(\mu) + C(\psi_i(z_0) - \psi_i(z))) p(d\mu, dz) = 0, \quad i = 1, \dots, N\}, \quad (40)$$

(compare with (8) and (10)). Note that

$$\tilde{a}_{M,N}(z_0) \leq \tilde{a}^*(z_0) \quad \forall M, N = 1, 2, \dots, \quad \lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \tilde{a}_{M,N}(z_0) = \tilde{a}^*(z_0) \quad (41)$$

(see Proposition 7.1 in [30]). Denote by $\zeta^{M,N}(z)$ a solution of the problem (37) (assuming that it exists) and consider the problem

$$\sup_{\eta(\cdot) \in \mathcal{V}_M} \min_{(u,y) \in U \times Y_z} \{r(u, y) + \nabla \zeta^{M,N}(z)^T h(u, y) + \nabla \eta(y)^T f(u, y)\} = \tilde{H}_M(\nabla \zeta^{M,N}(z), z) \quad \forall z \in Z, \quad (42)$$

where \sup is over the functions from the finite dimensional space $\mathcal{V}_M \subset C^1(\mathbb{R}^m)$,

$$\mathcal{V}_M \stackrel{\text{def}}{=} \{\eta(\cdot) \in C^1(\mathbb{R}^m) : \eta(y) = \sum_{i=1}^M \omega_i \phi_i(y), \quad \omega = (\omega_i) \in \mathbb{R}^M\}. \quad (43)$$

The problem (43) is dual with respect to the problem in the right-hand-side of (36) taken with $p = \nabla \zeta^{M,N}(z)$. Denote by $\eta_z^{M,N}(y)$ a solution of the problem (42) (assuming that it exists too) and define $u^{M,N}(y, z)$ as a minimizer

$$u^{M,N}(y, z) \stackrel{\text{def}}{=} \operatorname{argmin}_{u \in U} \{r(u, y) + \nabla \zeta^{M,N}(z)^T h(u, y) + \nabla \eta_z^{M,N}(y)^T f(u, y)\}, \quad (44)$$

For the class of problems considered in [30], it has been shown that, under certain conditions, the feedback control $u^{M,N}(y, z)$ generates a $\nu(M, N)$ -near optimal ACG family, with $\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \nu(M, N) = 0$ (see Theorem 9.4 in [30]). The argument used in [30] can be extended to the class of problems that we are dealing with. We, however, do not present this argument in the paper, and we conclude by just indicating a way to numerically evaluate a “degree of near optimality” of the control $u^{M,N}(y, z)$ defined in (44) (provided that it generates an ACG family). Note first of all that, if solutions of the problems (37) and (42) exist, then they are presentable in the form

$$\zeta^{M,N}(z) = \sum_{i=1}^N \lambda_i^{M,N} \psi_i(z), \quad \eta_z^{N,M}(y) = \sum_{i=1}^M \omega_{z,i}^{N,M} \phi_i(y). \quad (45)$$

The coefficients of the expansions in (45) can be numerically found via solving the semi-infinite LP problem (39) (see the description of the algorithm in Section 4.3 of [30]). Once these coefficients are found, one can define $u^{M,N}(y, z)$ according to (44) and verify whether the system

$$\frac{dy(\tau)}{d\tau} = f(u^{M,N}(y(\tau), z), y(\tau)), \quad y(0) = y_0 \in Y_z \quad (46)$$

has a solution $y_z^{M,N}(\tau)$ such that the pair $(u_z^{M,N}(\cdot), y_z^{M,N}(\cdot))$, where $u_z^{M,N}(\tau) = u^{M,N}(y_z^{M,N}(\tau), z)$, generates an occupational measure $\mu_z^{M,N}$. Finally, one can numerically integrate the system

$$z'(t) = \tilde{h}(\mu_{z(t)}^{M,N}), \quad z(0) = z_0, \quad (47)$$

and (provided that the solution $z^{M,N}(t)$ of the latter exists and is contained in Z) evaluate the integral

$$\int_0^{+\infty} e^{-Ct} \tilde{r}(\mu_{z^{M,N}(t)}^{M,N}) dt \stackrel{\text{def}}{=} \tilde{R}^{M,N}(y_0), \quad (48)$$

the latter being the value of the objective function of the averaged system (8) obtained with $\mu(t) = \mu_{z^{M,N}(t)}^{M,N}$. By (11) and (41),

$$\tilde{R}^{M,N}(y_0) - R^*(y_0) \leq \tilde{R}^{M,N}(y_0) - \frac{\tilde{a}_{M,N}(z_0)}{C}, \quad (49)$$

and, thus, the right-hand-side in (49) gives an estimate of near optimality of the ACG family generated by the control $u^{M,M}(y, z)$.

6 Construction of asymptotically optimal/near optimal controls for the perturbed problem

In this section, we will describe a way how asymptotically optimal or near optimal controls for the perturbed problem can be constructed on the basis of controls that generate optimal/near optimal ACG families. Everywhere in what follows it is assumed that all the conditions of Corollary 2 are satisfied and, hence, (30) is valid (the limit of the optimal value of the perturbed problem exists and is equal to the optimal value of the averaged problem). Moreover, for simplicity of the exposition, we restrict ourselves to the consideration of the case when the

reduced system is uncontrolled (that is, it is of the form (1)) and we assume that Assumption 3 is valid. Note that the perturbed system (11) takes the form

$$\frac{dy_\epsilon(t)}{dt} = \frac{1}{\epsilon} f(y_\epsilon(t)) + g(u(t), y_\epsilon(t)), \quad y(0) = y_0 \quad (1)$$

in this case.

Definition 4 A control $u_\epsilon^*(\cdot)$ is called asymptotically optimal for the perturbed problem (10) if

$$\lim_{\epsilon \rightarrow 0} \int_0^{+\infty} e^{-Ct} r(u_\epsilon^*(t), y_\epsilon^*(t)) dt = \lim_{\epsilon \rightarrow 0} R^*(\epsilon, y_0), \quad (2)$$

where $y_\epsilon^*(\cdot)$ is the solution of the system (11) obtained with the control $u_\epsilon^*(\cdot)$. A control $u_\epsilon^*(\cdot)$ is called asymptotically α -near optimal ($\alpha > 0$) for the perturbed problem (10) if

$$\lim_{\epsilon \rightarrow 0} \int_0^{+\infty} e^{-Ct} r(u_\epsilon^*(t), y_\epsilon^*(t)) dt \leq \lim_{\epsilon \rightarrow 0} R^*(\epsilon, y_0) + \alpha. \quad (3)$$

Let $u^*(y, z)$ be a feedback control that generates an optimal or near optimal ACG family and let μ_z^* be the occupational measure that is generated by the pair $(u^*(y_z(\tau), z), y_z(\tau))$, where $y_z(\cdot)$ is a solution of (1). The latter is T_z -periodic (due to Assumption 3) and, hence, for any continuous function $q(u, y)$,

$$\int_{Y_z} q(u, y) \mu_z^*(du, dy) = \frac{1}{T_z} \int_0^{T_z} q(u^*(y_z(\tau), z), y_z(\tau)) d\tau. \quad (4)$$

Let us construct a control $u_\epsilon^*(t)$ that will be shown to be asymptotically optimal (near optimal) in the perturbed problem. To this end, partition the interval $[0, \infty)$ by the points

$$t_l = l\Delta(\epsilon), \quad l = 0, 1, \dots \quad (5)$$

where

$$\Delta(\epsilon) = \frac{\epsilon}{2L_f} \ln \frac{1}{\epsilon}, \quad (6)$$

L_f being a Lipschitz constant of $f(\cdot)$. Note that

$$\lim_{\epsilon \rightarrow 0} \Delta(\epsilon) = 0, \quad \lim_{\epsilon \rightarrow 0} \frac{\Delta(\epsilon)}{\epsilon} = \infty. \quad (7)$$

Define $u_\epsilon^*(t)$ on the interval $[0, t_1)$ by the equation

$$u_\epsilon^*(t) = u^*(y_{y_0}(\frac{t}{\epsilon}), z_0) \quad \forall t \in [0, t_1), \quad (8)$$

where $y_{y_0}(\tau)$ is the solution of the system (1) considered on the interval $[0, \frac{t_1}{\epsilon}]$ with the initial condition $y_{y_0}(0) = y_0$ (recall that $z_0 \stackrel{\text{def}}{=} F(y_0)$).

Let us assume that the control $u_\epsilon^*(t)$ has been defined on the interval $t \in [0, t_l)$ and let $y_\epsilon^*(t)$ be the corresponding solution of the perturbed system (1) on this interval. Extend the definition of $u_\epsilon^*(t)$ to the interval $[0, t_{l+1}]$ by taking

$$u_\epsilon^*(t) = u^*(y_{y_\epsilon^*(t_l)}(\frac{t}{\epsilon}), z_\epsilon^*(t_l)) \quad \forall t \in [t_l, t_{l+1}), \quad l = 1, 2, \dots, \quad (9)$$

where $y_{y_\epsilon^*(t_l)}(\tau)$ is the solution of the system (1) considered on the interval $[\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]$ with the initial condition $y_{y_\epsilon^*(t_l)}(\frac{t_l}{\epsilon}) = y_\epsilon^*(t_l)$, and where $z_\epsilon^*(t_l) \stackrel{\text{def}}{=} F(y_\epsilon^*(t_l))$.

Proposition 10 *Let $u^*(y, z)$ generate an optimal (α -near optimal) ACG family. Assume that*

$$y_\epsilon^*(t) \in Y, \quad F(y_\epsilon^*(t)) \in Z \quad \forall t \in [0, \infty), \quad (10)$$

where $y_\epsilon^(\cdot)$ is the solution of the system (1) obtained with use of the control $u_\epsilon^*(\cdot)$ constructed as described above. Assume also that the function $\tilde{h}^*(z)$,*

$$\tilde{h}^*(z) \stackrel{\text{def}}{=} \int_{Y_z} h(u, y) \mu_z^*(du, dy) = \frac{1}{T_z} \int_0^{T_z} h(u^*(y_z(\tau), z), y_z(\tau)) d\tau \quad (11)$$

and the function $\tilde{r}^(z)$,*

$$\tilde{r}^*(z) \stackrel{\text{def}}{=} \int_{Y_z} r(u, y) \mu_z^*(du, dy) = \frac{1}{T_z} \int_0^{T_z} r(u^*(y_z(\tau), z), y_z(\tau)) d\tau \quad (12)$$

satisfy Lipschitz conditions in a neighbourhood of the solution $z^(\cdot)$ of the system*

$$z'(t) = \tilde{h}^*(z(t)), \quad z(0) = z_0. \quad (13)$$

Then the control $u_\epsilon^(t)$ is asymptotically optimal (α -near optimal) in the perturbed problem (10).*

Proof. The proof of the proposition is given in Section 7. \square

The construction of an asymptotically optimal (near optimal) control is simplified if the function $u^*(y, z)$ is continuous.

Proposition 11 *Let $u^*(y, z)$ that generate an optimal (α -near optimal) ACG family be continuous function of (y, z) and let the solution $y_\epsilon^*(\cdot)$ of the system*

$$\frac{dy_\epsilon^*(t)}{dt} = \frac{1}{\epsilon} f(y_\epsilon^*(t)) + g(u^*(y_\epsilon^*(t), F(y_\epsilon^*(t))), y_\epsilon^*(t)), \quad y_\epsilon^*(0) = y_0 \quad (14)$$

satisfy the inclusions (10). Let the functions $\tilde{h}^(z)$, $\tilde{r}^*(z)$ be Lipschitz continuous in a neighbourhood of the solution $z^*(\cdot)$ of the system (13). Then the feedback control $u^*(y, F(y))$ is asymptotically optimal (α -near optimal) in (10). That is,*

$$\lim_{\epsilon \rightarrow 0} \int_0^{+\infty} e^{-Ct} r(u^*(y_\epsilon^*(t), F(y_\epsilon^*(t))), y_\epsilon^*(t)) dt = \lim_{\epsilon \rightarrow 0} R^*(\epsilon, y_0) \quad (\leq \lim_{\epsilon \rightarrow 0} R^*(\epsilon, y_0) + \alpha). \quad (15)$$

Proof. The proof is in Section 7. \square

Remark 3 Note that the assumption that the functions $\tilde{h}^*(z)$ and $\tilde{r}^*(z)$ are Lipschitz continuous in a neighbourhood of $z^*(\cdot)$ can be replaced by a weaker assumption that they are piece-wise Lipschitz continuous in a neighbourhood of $z^*(\cdot)$ (see Definition 4.4 in [30]). The proof of the result is, however, much more technical in this case and we do not present it in the paper.

Remark 4 Assume that the conditions of Proposition 11 are satisfied with $u^*(y, z) = u^{M,N}(y, z)$, where $u^{M,N}(y, z)$ is defined as a minimizer (44). Denote by $R^{M,N}(\epsilon, y_0)$ the value of the objective function (10) obtained with the use of the feedback control $u^{M,N}(y, F(y))$ in the perturbed system. Due to (30) and due to (11), (41),

$$\lim_{\epsilon \rightarrow 0} R^{M,N}(\epsilon, y_0) - \lim_{\epsilon \rightarrow 0} R^*(\epsilon, y_0) = \lim_{\epsilon \rightarrow 0} R^{M,N}(\epsilon, y_0) - \tilde{R}^*(y_0) \leq \lim_{\epsilon \rightarrow 0} R^{M,N}(\epsilon, y_0) - \frac{\tilde{a}_{M,N}(z_0)}{C}$$

(compare with (49)). Thus, $R^{M,N}(\epsilon, y_0) - \frac{\tilde{a}_{M,N}(z_0)}{C}$ may serve as a measure of near optimality of the feedback control $u^{M,N}(y, F(y))$ in the perturbed problem (provided that ϵ is small enough).

Let us demonstrate the construction of asymptotically near optimal control in the problem of minimization considered on the solutions of the system (17) (see Example 2 in Section 4)

Example 2 (continued). Consider the problem of optimal control

$$\epsilon \inf_{u(\cdot)} \int_0^\infty e^{-0.1\epsilon\tau} [u^2(\tau) + (F(y(\tau)) + 2.05)^2] d\tau \stackrel{\text{def}}{=} R^*(\epsilon, y_0) \quad (16)$$

where *inf* is over all controls and the corresponding solutions of the system (17) that satisfy the initial condition $y(0) = y_0 = (0.8916, 3.1370)$ (note that $z_0 \stackrel{\text{def}}{=} F(y_0) = -3$). Differentiating the observable $z(\tau) = F(y(\tau))$ (with $F(y)$ being as in (19)) along a solution of the system (17), one can readily come to the conclusion that its dynamics is described by the equation (compare with the general case (5))

$$\frac{dz(\tau)}{d\tau} = \epsilon u(\tau)(y_1(\tau) - 1), \quad z(0) = z_0. \quad (17)$$

Note that, the solutions of (17) and (17) are written without the subscript ϵ (to simplify the notations). Having in mind the structure of the objective function (16), one can expect that the control minimizing (16) will make the state trajectory spiral down from a larger orbit corresponding to $z_0 = -3$ to a smaller orbit corresponding to $z_0 = -2.05$ and that the solution of the augmented perturbed system (17), (17) obtained with this control stays in $Y \times Z$ (where Y and Z are defined in (20)). This kind of behaviour was indeed obtained with the use of control constructed as described at the end of Section 5.

In more detail, the problems (37), (39) and (42) were formed with the use of the monomials $\psi(z) = z^i$, $i = 1, \dots, 10$, and $\phi(y) = y_1^{i_1} y_2^{i_2}$, $1 \leq i_1 + i_2 \leq 5$. That is, $N = 10$ (the dimension of the space \mathcal{Q}_N in (38)) and $M = 35$ (the dimension of the space \mathcal{V}_M in (43)). The coefficients of the expansions (45) for $\zeta^{M,N}(z)$ and $\eta_z^{M,N}(y)$ were numerically found with the help of a software implementation of the algorithm described in Section 4.3 of [30]. Note that the derivative of the function $\zeta^{M,N}(z)$ was verified to be negative

$$\frac{d\zeta^{M,N}(z)}{dz} < 0 \quad \forall z \in [-3, -2.05]. \quad (18)$$

Having in mind the fact that in the given example

$$f(u, y) = f(y), \quad h(u, y) = u(y_1 - 1), \quad r(u, y) = u^2 + (F(y) + 2.05)^2, \quad U = [0, 1],$$

one can readily establish that the minimizer in (44) is of the form

$$u^{M,N}(y, z) = \min\left\{-\frac{1}{2} \frac{d\zeta^{M,N}(z)}{dz} (y_1 - 1), 1\right\} \quad \text{if } y_1 > 1; \quad u(y, z) = 0 \quad \text{if } y_1 \leq 1. \quad (19)$$

Note that, due to (18), $u^{M,N}(y, z)$ is positive if $y_1 > 1$ and it is zero if $y_1 \leq 1$. Applying the feedback control $u^{M,N}(y, F(y))$ in system (17) (considered with $\epsilon = 0.1$) until the moment the state trajectory reaches the smaller orbit characterized by the equality $F(y) = -2.05$ and applying the “zero control” after that moment, one obtains the solution $y(t) = (y_1(t), y_2(t))$. The state trajectory of this solution and the graph of the control $u(t) = u^{M,N}(y(t), F(y(t)))$ are depicted in Figures 2 and 3. The graph of $y_1(t)$ is depicted in figure 4 (the graph of $y_2(t)$ looks similar). The value of the objective function (16) thus obtained is ≈ 1.25 ($R^{M,N}(\epsilon, y_0) \approx 1.25$ with $\epsilon = 0.1$). The optimal value of the objective function of the semi-infinite LP problem (39) was numerically evaluated to be ≈ 0.117 ($\tilde{a}_{M,N}(z_0) \approx 0.117$). That is, $\frac{\tilde{a}_{M,N}(z_0)}{C} \approx 1.17$ as $C = 0.1$ in this case. Thus, the difference $1.25 - 1.17 = 0.08$ provides us with a measure of near optimality of the found solution (see Remark 4).

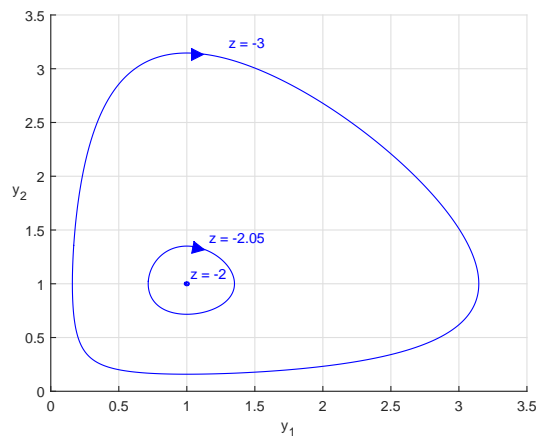


Fig. 1: State trajectories for $\epsilon = 0$

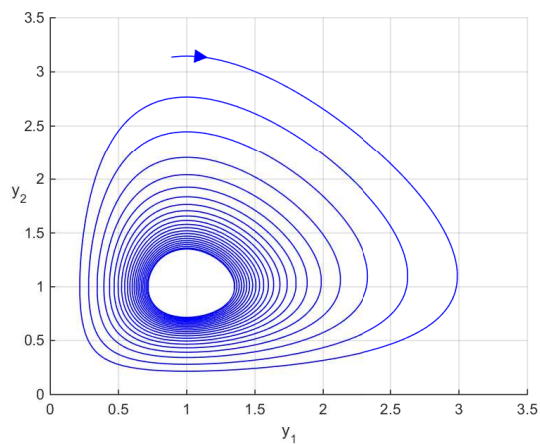


Fig. 2: Near optimal state trajectory for $\epsilon = 0.1$

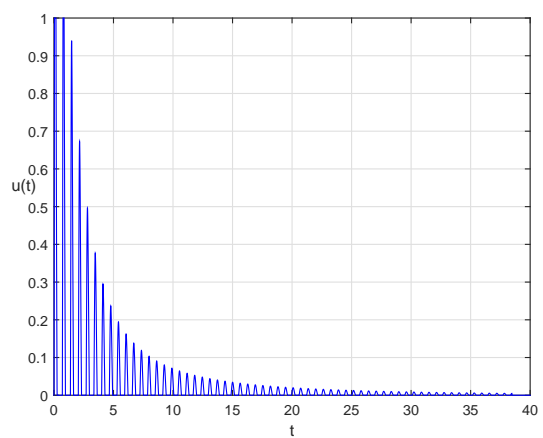


Fig. 3: $u(t) = u^{M,N}(y(t), F(y(t)))$

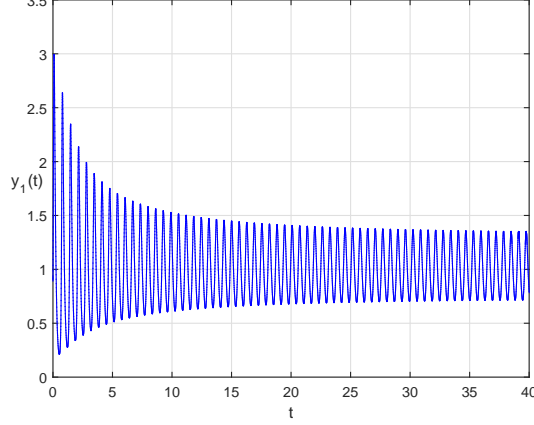


Fig. 4: $y_1(t)$

7 Proofs of Theorems 1 and 2 and a proof of Proposition 10

Proof of Theorem 1. Let

$$X(t, u, y) \stackrel{\text{def}}{=} (h(u, y), e^{-Ct}r(u, y)) \quad (1)$$

and let

$$V(t, z) \stackrel{\text{def}}{=} \cup_{\mu \in W(z)} \left\{ \int_{U \times Y} X(t, u, y) \mu(du, dy) \right\} = \cup_{\mu \in W(z)} \{(\tilde{h}(\mu), e^{-Ct}\tilde{r}(\mu))\}. \quad (2)$$

Note that the map $V(t, z)$ is convex and compact valued (due to the fact that $W(z) : Z \rightsquigarrow \mathcal{P}(U \times Y)$ is convex and compact valued). Also from (15) it follows that

$$d_H(V(t', z'), V(t'', z'')) \leq \hat{L}(|t' - t''| + \|z' - z''\|) \quad \forall t', t'' \geq 0, \forall z', z'' \in Z, \quad \hat{L} = \text{const.} \quad (3)$$

Let $\theta_\epsilon(t)$ stand for the solution of the equation

$$\frac{d\theta_\epsilon(t)}{dt} = e^{-Ct}r(u_\epsilon(t), y_\epsilon(t)), \quad \theta_\epsilon(0) = 0. \quad (4)$$

Then $x_\epsilon(t) \stackrel{\text{def}}{=} (z_\epsilon(t), \theta_\epsilon(t))$ is the solution of the equation (from this point on, we will write $u(t)$ instead of $u_\epsilon(t)$ to simplify the notations)

$$\frac{dx_\epsilon(t)}{dt} = X(t, u(t), y_\epsilon(t)), \quad x_\epsilon(0) = (z_0, 0). \quad (5)$$

Due to Filippov's selection theorem (see, e.g., Theorem 8.2.10 in [9]), to prove the estimates (22) and (23), it is sufficient to show that there exists a solution $x(t) = (z(t), \theta(t))$ of the differential inclusion

$$\frac{dx(t)}{dt} \in V(t, z(t)), \quad x(0) = (z_0, 0) \quad (6)$$

such that

$$\max_{t \in [0, T]} \|x_\epsilon(t) - x(t)\| \leq \beta(\epsilon, T). \quad (7)$$

From the validity of (22) it will follow that $z(t) \in Z \forall t \in [0, T]$ (for ϵ small enough). Also the solution $x(t)$ of (7) is extendable to the interval $[T, \infty)$ in such a way that $z(t) \in Z \forall t \in [T, \infty)$ (due to the viability of the averaged system). Thus, the theorem will be proved if one establishes that a solution of (6) satisfying (7) exists.

Since $x_\epsilon(t)$ satisfies (5), one can write down the following relationships

$$x_\epsilon(t_{l+1}) = x_\epsilon(t_l) + \int_{t_l}^{t_{l+1}} X(t, u(t), y_\epsilon(t)) dt = x_\epsilon(t_l) + \epsilon \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(\epsilon\tau, u(\epsilon\tau), y_\epsilon(\epsilon\tau)) d\tau, \quad l = 0, 1, \dots, N_\epsilon - 1, \quad (8)$$

where t_l are as defined in (5) and $N_\epsilon \stackrel{\text{def}}{=} \lfloor \frac{T}{\Delta(\epsilon)} \rfloor$ ($\lfloor \cdot \rfloor$ standing for the floor function). Consider also the difference equation

$$\bar{x}_\epsilon(l+1) = \bar{x}_\epsilon(l) + \Delta(\epsilon) \int_{U \times Y} X(t_l, u, y) \bar{\mu}_l(du, dy), \quad l = 0, 1, \dots, N_\epsilon - 1, \quad \bar{x}_\epsilon(0) = (z_0, 0), \quad (9)$$

where $\bar{x}_\epsilon(l) \stackrel{\text{def}}{=} (\bar{z}_\epsilon(l), \bar{\theta}_\epsilon(l))$, and $\bar{\mu}_l(du, dy)$ are constructed iteratively in such a way that

$$\bar{\mu}_l \in W(\bar{z}_\epsilon(l)) \quad (10)$$

and

$$\left| \left(\frac{\Delta(\epsilon)}{\epsilon} \right)^{-1} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u(\epsilon\tau), y_\epsilon(\epsilon\tau)) d\tau - \int_{U \times Y} X(t_l, u, y) \bar{\mu}_l(du, dy) \right| \leq \hat{L} \|z_\epsilon(t_l) - \bar{z}_\epsilon(l)\| + \beta_1(\epsilon) \quad (11)$$

for any $l = 0, 1, \dots, N_\epsilon - 1$ (with \hat{L} being a Lipschitz constant from (3) and with $\lim_{\epsilon \rightarrow 0} \beta_1(\epsilon) = 0$; the construction of $\bar{\mu}_l$, $l = 0, 1, \dots, N_\epsilon - 1$, that satisfy (10) and (11) is described at the end of the proof of the theorem).

By subtracting (9) from (8), one obtains the inequalities

$$\begin{aligned} \|x_\epsilon(t_{l+1}) - \bar{x}_\epsilon(l+1)\| &\leq \|x_\epsilon(t_l) - \bar{x}_\epsilon(l)\| + \Delta(\epsilon)(\hat{L}\|z_\epsilon(t_l) - \bar{z}_\epsilon(l)\| + \beta_2(\epsilon)) \\ &\leq \|x_\epsilon(t_l) - \bar{x}_\epsilon(l)\| + \Delta(\epsilon)(\hat{L}\|x_\epsilon(t_l) - \bar{x}_\epsilon(l)\| + \beta_2(\epsilon)), \quad l = 0, 1, \dots, N_\epsilon - 1, \end{aligned} \quad (12)$$

where $\lim_{\epsilon \rightarrow 0} \beta_2(\epsilon) = 0$. These imply (see Proposition 5.1 in [22]) that

$$\|x_\epsilon(t_l) - \bar{x}_\epsilon(l)\| \leq \kappa(\epsilon, T), \quad l = 0, 1, \dots, N_\epsilon - 1, \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \kappa(\epsilon, T) = 0. \quad (13)$$

Define the piecewise linear function $\bar{\bar{x}}_\epsilon(t) = (\bar{\bar{z}}_\epsilon(t), \bar{\bar{\theta}}_\epsilon(t))$ on the interval $[0, T]$ by the equations

$$\begin{aligned} \bar{\bar{x}}_\epsilon(t) &= \bar{x}_\epsilon(l) + (t - t_l)v_l \quad \forall t \in [t_l, t_{l+1}), \quad l = 0, 1, \dots, N_\epsilon - 1, \\ \bar{\bar{x}}_\epsilon(t) &= \bar{x}_\epsilon(l) + (t - t_{N_\epsilon})v_{N_\epsilon-1} \quad \forall t \in [t_{N_\epsilon}, t_{N_\epsilon+1}], \quad t_{N_\epsilon+1} \stackrel{\text{def}}{=} T, \end{aligned}$$

where

$$v_l \stackrel{\text{def}}{=} \int_{U \times Y} X(t_l, u, y) \bar{\mu}_l(du, dy) \in V(t_l, \bar{z}_\epsilon(l)).$$

Note that

$$\max_{t \in [t_l, t_{l+1}]} \|\bar{\bar{x}}_\epsilon(t) - \bar{x}_\epsilon(l)\| \leq \Delta(\epsilon) \|v_l\| \leq M \Delta(\epsilon), \quad \text{where} \quad M \stackrel{\text{def}}{=} \max_{(u, y) \in U \times Y} (|h(u, y)| + |r(u, y)|). \quad (14)$$

Also, due to (3), for $t \in (t_l, t_{l+1})$ (with $l = 0, 1, \dots, N_\epsilon$),

$$\begin{aligned} d\left(\frac{d\bar{\bar{x}}_\epsilon(t)}{dt}, V(t, \bar{\bar{z}}_\epsilon(t))\right) &\leq d(v_l, V(t_l, \bar{z}_\epsilon(l))) + d_H(V(t_l, \bar{z}_\epsilon(l)), V(t, \bar{\bar{z}}_\epsilon(t))) \\ &= d_H(V(t_l, \bar{z}_\epsilon(l)), V(t, \bar{\bar{z}}_\epsilon(t))) \leq L_1 \Delta(\epsilon), \quad L_1 = \text{const.} \end{aligned}$$

From Filippov existence theorem (see, e.g. Theorem 10.4.1 in [9]) it now follows that there exists a solution $x(t)$ of the differential inclusion (6) such that

$$\max_{t \in [0, T]} \|\bar{\bar{x}}_\epsilon(t) - x(t)\| \leq L_T \Delta(\epsilon), \quad L_T = \text{const.} \quad (15)$$

This (along with (13) and (14)) prove the required statement.

To finalize the proof of the theorem, let us now describe the construction of $\bar{\mu}_l$ ($l = 0, 1, \dots, N_\epsilon - 1$) satisfying (10) and (11). For $\tau \in [\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]$, let $y_{y_\epsilon(t_l)}(\tau)$ be the solution of the system (1) that is obtained with the use of the control $u(\epsilon\tau)$ and that satisfies the initial condition $y_{y_\epsilon(t_l)}(\frac{t_l}{\epsilon}) = y_\epsilon(t_l)$. That is,

$$y_{y_\epsilon(t_l)}(\tau) = y_\epsilon(t_l) + \int_{\frac{t_l}{\epsilon}}^{\tau} f(u(\epsilon\tau'), y_{y_\epsilon(t_l)}(\tau')) d\tau' \quad \forall \tau \in [\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]. \quad (16)$$

Due to the fact that $y_\epsilon(t)$ is a solution of (11), it satisfies the equation

$$y_\epsilon(\epsilon\tau) = y_\epsilon(t_l) + \int_{\frac{t_l}{\epsilon}}^{\tau} [f(u(\epsilon\tau'), y_\epsilon(\epsilon\tau')) + \epsilon g(u(\epsilon\tau'), y_\epsilon(\epsilon\tau'))] d\tau' \quad \forall \tau \in [\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]. \quad (17)$$

Hence,

$$\|y_\epsilon(\epsilon\tau) - y_{y_\epsilon(t_l)}(\tau)\| \leq L_f \int_{\frac{t_l}{\epsilon}}^{\tau} \|y_\epsilon(\epsilon\tau') - y_{y_\epsilon(t_l)}(\tau')\| d\tau' + M_g \Delta(\epsilon) \quad \forall \tau \in [\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}], \quad (18)$$

where L_f is a Lipschitz constant of $f(\cdot)$ and $M_g \stackrel{\text{def}}{=} \max_{(u,y) \in U \times Y} \|g(u,y)\|$. By Gronwall-Bellman lemma, the latter implies that

$$\max_{\tau \in [\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]} \|y_\epsilon(\epsilon\tau) - y_{y_\epsilon(t_l)}(\tau)\| \leq M_g \Delta(\epsilon) e^{L_f \frac{\Delta(\epsilon)}{\epsilon}} \leq M_g \left(\frac{\epsilon}{2L_f} \ln \frac{1}{\epsilon} \right) \left(\frac{1}{\epsilon^{\frac{1}{2}}} \right) < \epsilon^{\frac{1}{4}}, \quad (19)$$

the second to last inequality being valid due to (6) (ϵ is assumed to be small enough). Note that from (19) it follows that

$$\|(\frac{\Delta(\epsilon)}{\epsilon})^{-1} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u(\epsilon\tau), y_\epsilon(\epsilon\tau)) d\tau - (\frac{\Delta(\epsilon)}{\epsilon})^{-1} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u(\epsilon\tau), y_{y_\epsilon(t_l)}(\tau)) d\tau\| \leq \hat{L} \epsilon^{\frac{1}{4}}. \quad (20)$$

Denote by μ_l the occupational measure generated by the pair $(u(\epsilon\tau), y_{y_\epsilon(t_l)}(\tau))$ on the interval $[\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]$. That is, μ_l is such that for any continuous function $q(u, y)$

$$\int_{U \times Y} q(u, y) \mu_l(du, dy) = (\frac{\Delta(\epsilon)}{\epsilon})^{-1} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} q(u(\epsilon\tau), y_{y_\epsilon(t_l)}(\tau)) d\tau. \quad (21)$$

By Proposition 1 (see (16)), there exists $\tilde{\mu}_l \in W(z_\epsilon(t_l))$ such that

$$\rho(\mu_l, \tilde{\mu}_l) \leq \beta((\frac{\Delta(\epsilon)}{\epsilon})^{-1}) \stackrel{\text{def}}{=} \beta_3(\epsilon), \quad \lim_{\epsilon \rightarrow 0} \beta_3(\epsilon) = 0, \quad (22)$$

the latter implying that

$$\|(\frac{\Delta(\epsilon)}{\epsilon})^{-1} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u(\epsilon\tau), y_{y_\epsilon(t_l)}(\tau)) d\tau - \int_{U \times Y} X(t_l, u, y) \tilde{\mu}_l(du, dy)\| \leq \beta_4(\epsilon), \quad \lim_{\epsilon \rightarrow 0} \beta_4(\epsilon) = 0. \quad (23)$$

Also, due to (3), there exists $\bar{\mu}_l \in W(\bar{z}_\epsilon(l))$ such that

$$\|\int_{U \times Y} X(t_l, u, y) \tilde{\mu}_l(du, dy) - \int_{U \times Y} X(t_l, u, y) \bar{\mu}_l(du, dy)\| \leq \hat{L} \|z_\epsilon(t_l) - \bar{z}_\epsilon(l)\|. \quad (24)$$

Summarizing (20), (23) and (24), one obtains that thus defined $\bar{\mu}_l$ satisfies (11) (with $\beta_1(\epsilon) = \hat{L} \epsilon^{\frac{1}{4}} + \beta_4(\epsilon)$). \square

Proof of Theorem 2. Let $(\mu(\cdot), z(\cdot))$ satisfy (6)-(7) and let $\theta(t)$ be the solution of the equation

$$\frac{d\theta(t)}{dt} = e^{-Ct}\tilde{r}(\mu(t)), \quad \theta(0) = 0. \quad (25)$$

Then $x(\cdot) = (z(t), \theta(t))$ is a solution of the differential inclusion (6) and, due to (3), for any $t \in [t_l, t_{l+1}]$ ($t_l, l = 0, \dots, N_\epsilon - 1$, are as in (5)),

$$\frac{dx(t)}{dt} \in V(t, z(t)) \subset V(t_l, z(t_l)) + \hat{L}(|t - t_l| + \|z(t) - z(t_l)\|)\bar{B} \subset V(t_l, z(t_l)) + L_2\Delta(\epsilon)\bar{B}, \quad (26)$$

where L_2 is a sufficiently large constant and \bar{B} is the closed unit ball in \mathbb{R}^{k+1} . Consequently,

$$\begin{aligned} & \Delta(\epsilon)^{-1} \int_{t_l}^{t_{l+1}} \frac{dx(t)}{dt} dt \in V(t_l, z(t_l)) + L_2\Delta(\epsilon)\bar{B}, \quad l = 0, \dots, N_\epsilon - 1 \\ \Rightarrow & \text{dist} \left(\Delta(\epsilon)^{-1} \int_{t_l}^{t_{l+1}} \frac{dx(t)}{dt} dt, V(t_l, z(t_l)) \right) \leq L_2\Delta(\epsilon), \quad l = 0, \dots, N_\epsilon - 1. \end{aligned} \quad (27)$$

Let $\bar{v}_l \in V(t_l, z(t_l))$, $l = 0, \dots, N_\epsilon - 1$, be defined by the equation

$$\bar{v}_l \stackrel{\text{def}}{=} \text{argmin} \{ \|v - \Delta(\epsilon)^{-1} \int_{t_l}^{t_{l+1}} \frac{dx(t)}{dt} dt\| \mid v \in V(t_l, z(t_l)) \}. \quad (28)$$

Note that, by (27),

$$\|\bar{v}_l - \Delta(\epsilon)^{-1} \int_{t_l}^{t_{l+1}} \frac{dx(t)}{dt} dt\| \leq L_2\Delta(\epsilon), \quad l = 0, \dots, N_\epsilon - 1, \quad (29)$$

and note that from the definition of the multivalued map $V(\cdot, \cdot)$ (see (2)) it follows that there exists $\bar{\mu}_l \in W(z(t_l))$ such that

$$\bar{v}_l = \int_{U \times Y} X(t_l, u, y) \bar{\mu}_l(du, dy), \quad l = 0, \dots, N_\epsilon - 1. \quad (30)$$

Consider the difference equation

$$\bar{x}_\epsilon(l+1) = \bar{x}_\epsilon(l) + \Delta(\epsilon)\bar{v}_l, \quad l = 0, 1, \dots, N_\epsilon - 1, \quad \bar{x}_\epsilon(0) = (z_0, 0), \quad (31)$$

where $\bar{x}_\epsilon(l) \stackrel{\text{def}}{=} (\bar{z}_\epsilon(l), \bar{\theta}_\epsilon(l))$. Subtracting the latter from

$$x(t_{l+1}) = x(t_l) + \Delta(\epsilon) \left(\Delta(\epsilon)^{-1} \int_{t_l}^{t_{l+1}} \frac{dx(t)}{dt} dt \right), \quad l = 0, 1, \dots, N_\epsilon - 1, \quad \bar{z}_\epsilon(0) = z_0, \quad (32)$$

one can establish the validity of the inequalities

$$\|x(t_{l+1}) - \bar{x}_\epsilon(l+1)\| \leq \|x(t_l) - \bar{x}_\epsilon(l)\| + L_2\Delta^2(\epsilon), \quad l = 0, 1, \dots, N_\epsilon - 1. \quad (33)$$

These imply (see Proposition 5.1 in [22])

$$\|x(t_l) - \bar{x}_\epsilon(l)\| \leq \kappa_1(\epsilon, T), \quad l = 0, 1, \dots, N_\epsilon - 1, \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \kappa_1(\epsilon, T) = 0. \quad (34)$$

Assume now that the control $u(t)$ has been defined on the interval $[0, t_l]$, $l < N_\epsilon$, and extend its definition to the interval $[0, t_{l+1}]$. Denote by $(y_\epsilon(t), x_\epsilon(t))$ the solution of the system (11),(5) obtained with the use of the control $u(t)$ on the interval $[0, t_l]$. Note that, due to (3), there exists a measure $\tilde{\mu}_l \in W(z_\epsilon(t_l))$ such that (see (30))

$$\left\| \int_{U \times Y} X(t_l, u, y) \tilde{\mu}_l(du, dy) - \bar{v}_l \right\| = \left\| \int_{U \times Y} X(t_l, u, y) \tilde{\mu}_l(du, dy) - \int_{U \times Y} X(t_l, u, y) \bar{\mu}_l(du, dy) \right\|$$

$$\leq \hat{L} \|z_\epsilon(t_l) - z(t_l)\|. \quad (35)$$

Define the control $u(t)$ on the interval $[t_l, t_{l+1}]$ in such a way that the pair $(u(\epsilon\tau), y_{y_\epsilon(t_l)}(\tau))$ (where $y_{y_\epsilon(t_l)}(\tau)$ is the solution of the reduced system (1) on the interval $[\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]$ which is obtained with the use of the control $u(\epsilon\tau)$ and with the initial condition $y_{y_\epsilon(t_l)}(\frac{t_l}{\epsilon}) = y_\epsilon(t_l)$) generates an occupational measure μ_l satisfying (22) and (23) (the existence of such control follows from Assumption 2). Use the control $u(t)$ to obtain the solution $(y_\epsilon(t), x_\epsilon(t))$ of the system (11),(5) on the interval $[t_l, t_{l+1}]$. Note that, similarly to the proof of Theorem 1, it can be verified that the estimate (20) is valid, which along with (23) and (35) imply that

$$\|(\frac{\Delta(\epsilon)}{\epsilon})^{-1} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u(\epsilon\tau), y_\epsilon(\epsilon\tau)) d\tau - \bar{v}_l\| \leq \hat{L} \|z_\epsilon(t_l) - z(t_l)\| + \beta_1(\epsilon), \quad (36)$$

where $\lim_{\epsilon \rightarrow 0} \beta_1(\epsilon) = 0$.

By continuing this process, one can construct the control $u(t)$ on the interval $[0, t_{N_\epsilon}]$. On the interval $[t_{N_\epsilon}, T]$, the control $u(t)$ can be taken to be equal to an arbitrary $u \in U$. Let us show that the solution $(y_\epsilon(t), x_\epsilon(t))$ of the system (11),(5) obtained with this control satisfies (22) and (23).

Having in mind the validity of (36) and arguing as in the proof of Theorem 1, one can establish the validity of the estimates (13), where $\{\bar{x}_\epsilon(l)\}$ is the solution of (31). These and (34) imply that

$$\|x_\epsilon(t_l) - x(t_l)\| \leq \kappa_2(\epsilon, T), \quad l = 0, 1, \dots, N_\epsilon - 1, \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \kappa_2(\epsilon, T) = 0. \quad (37)$$

Since

$$\max_{t \in [t_l, t_{l+1}]} \|x_\epsilon(t) - x_\epsilon(t_l)\| \leq M\Delta(\epsilon), \quad \max_{t \in [t_l, t_{l+1}]} \|x(t) - x(t_l)\| \leq M\Delta(\epsilon),$$

the estimates (37) imply the validity of (7) (with $\beta(\epsilon, T) \stackrel{\text{def}}{=} \kappa_2(\epsilon, T) + 2M\Delta(\epsilon)$), which in turn implies (22) and (23). Due to (22) and (19) and due to (24), the inclusions (25) are valid for $t \in [0, T]$. Since the augmented perturbed system (11), (12) is viable in $Y \times Z$, the solution of the latter can be extended in such a way that the inclusions (25) will be satisfied for all $t \in [0, \infty)$. This completes the proof of the theorem. \square

Proof of Proposition 10. Let $y_\epsilon^*(t)$ be the solution of the system (1) obtained with the control $u_\epsilon^*(t)$ and let $x_\epsilon^*(t) \stackrel{\text{def}}{=} (z_\epsilon^*(t), \theta_\epsilon^*(t))$ be defined by the equation

$$\frac{dx_\epsilon^*(t)}{dt} = X(t, u_\epsilon^*(t), y_\epsilon^*(t)), \quad x_\epsilon^*(0) = (z_0, 0), \quad (38)$$

where $X(\cdot)$ is as in (1). Note that a part of (38) is the equation

$$\frac{d\theta_\epsilon^*(t)}{dt} = e^{-ct} r(u_\epsilon^*(t), y_\epsilon^*(t)), \quad \theta_\epsilon^*(0) = 0. \quad (39)$$

From (38) it follows that

$$\begin{aligned} x_\epsilon^*(t_{l+1}) &= x_\epsilon^*(t_l) + \int_{t_l}^{t_{l+1}} X(t, u_\epsilon^*(t), y_\epsilon^*(t)) dt = x_\epsilon^*(t_l) + \int_{t_l}^{t_{l+1}} X(t, u^*(y_{y_\epsilon^*(t_l)}(\frac{t}{\epsilon}), z_\epsilon^*(t_l)), y_\epsilon^*(t)) dt \\ &= x_\epsilon^*(t_l) + \epsilon \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(\epsilon\tau, u^*(y_{y_\epsilon^*(t_l)}(\tau), z_\epsilon^*(t_l)), y_\epsilon^*(\epsilon\tau)) d\tau. \end{aligned} \quad (40)$$

Similarly to (19), one can verify that

$$\max_{\tau \in [\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]} \|y_\epsilon^*(\epsilon\tau) - y_{y_\epsilon^*(t_l)}(\tau)\| < \epsilon^{\frac{1}{4}}. \quad (41)$$

Consequently,

$$\begin{aligned} & \left\| \frac{\epsilon}{\Delta(\epsilon)} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u^*(y_{y_\epsilon^*}(t_l)(\tau), z_\epsilon^*(t_l)), y_\epsilon^*(\epsilon\tau)) d\tau - \frac{\epsilon}{\Delta(\epsilon)} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u^*(y_{y_\epsilon^*}(t_l)(\tau), z_\epsilon^*(t_l)), y_{y_\epsilon^*}(t_l)(\tau)) d\tau \right\| \\ & \leq L_X \epsilon^{\frac{1}{4}}, \end{aligned} \quad (42)$$

where L_X is a Lipschitz constant of $X(t, u, y)$ in y . Also, due to periodicity of $y_{y_\epsilon^*}(t_l)(\tau)$, one has (see (11))

$$\left\| \frac{\epsilon}{\Delta(\epsilon)} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u^*(y_{y_\epsilon^*}(t_l)(\tau), z_\epsilon^*(t_l)), y_{y_\epsilon^*}(t_l)(\tau)) d\tau - \tilde{X}(t_l, z_\epsilon^*(t_l)) \right\| \leq \bar{L} \frac{\epsilon}{\Delta(\epsilon)}, \quad (43)$$

where \bar{L} is a positive constant and where

$$\tilde{X}(t, z) \stackrel{\text{def}}{=} \frac{1}{T_z} \int_0^{T_z} X(t, u^*(y_z(\tau), z), y_z(\tau)) d\tau. \quad (44)$$

Note that, by (11) and (12),

$$\tilde{X}(t, z) = (\tilde{h}^*(z), e^{-Ct} \tilde{r}^*(z)). \quad (45)$$

Let $\theta^*(t)$ be the solution of the equation

$$\frac{d\theta(t)}{dt} = e^{-Ct} \tilde{r}^*(z^*(t)), \quad \theta(0) = 0 \quad (46)$$

and let $x(t) \stackrel{\text{def}}{=} (z^*(t), \theta^*(t))$, where $z^*(t)$ is the solution of (13). Then

$$\frac{dx(t)}{dt} = \tilde{X}(t, z^*(t)), \quad x(0) = (z_0, 0). \quad (47)$$

Consider the difference equation

$$x_\epsilon(l+1) = x_\epsilon(l) + \Delta(\epsilon) \tilde{X}(t_l, z_\epsilon(l)), \quad x_\epsilon(0) = (z_0, 0), \quad (48)$$

where $x_\epsilon(l) \stackrel{\text{def}}{=} (z_\epsilon(l), \theta_\epsilon(l))$. From the assumed Lipschitz continuity of $\tilde{h}(z)$ and $\tilde{r}(z)$ it follows that $\tilde{X}(t, z)$ is Lipschitz continuous, which, as can be readily verified, implies the validity of the estimates

$$\|x(t_l) - x_\epsilon(l)\| \leq \sigma_1(\epsilon, T), \quad l = 0, 1, \dots, \lfloor \frac{T}{\Delta(\epsilon)} \rfloor, \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \sigma_1(\epsilon, T) = 0, \quad (49)$$

where $x(t)$ is the solution of (47). By subtracting (48) from (40) and taking into account (42) and (43), one can obtain

$$\begin{aligned} \|x_\epsilon^*(t_{l+1}) - x_\epsilon(l+1)\| & \leq \|x_\epsilon^*(t_l) - x_\epsilon(l)\| + \Delta(\epsilon) \|\tilde{X}(t_l, z_\epsilon^*(t_l)) - \tilde{X}(t_l, z_\epsilon(l))\| + \bar{L}_1 \Delta(\epsilon) \left(\epsilon^{\frac{1}{4}} + \frac{\epsilon}{\Delta(\epsilon)} + \Delta(\epsilon) \right) \\ & \leq \|x_\epsilon^*(t_l) - x_\epsilon(l)\| + \Delta(\epsilon) L_{\tilde{X}} \|z_\epsilon^*(t_l) - z_\epsilon(l)\| + \bar{L}_1 \Delta(\epsilon) \left(\epsilon^{\frac{1}{4}} + \frac{\epsilon}{\Delta(\epsilon)} + \Delta(\epsilon) \right), \\ & \leq \|x_\epsilon^*(t_l) - x_\epsilon(l)\| + \Delta(\epsilon) L_{\tilde{X}} \|x_\epsilon^*(t_l) - x_\epsilon(l)\| + \bar{L}_1 \Delta(\epsilon) \left(\epsilon^{\frac{1}{4}} + \frac{\epsilon}{\Delta(\epsilon)} + \Delta(\epsilon) \right), \end{aligned}$$

where $L_{\tilde{X}}$ is a Lipschitz constant of $\tilde{X}(t, \cdot)$ and \bar{L}_2 is sufficiently large. The latter implies that

$$\|x_\epsilon^*(t_l) - x_\epsilon(l)\| \leq \sigma_2(\epsilon, T), \quad l = 0, 1, \dots, \lfloor \frac{T}{\Delta(\epsilon)} \rfloor, \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \sigma_2(\epsilon, T) = 0, \quad (50)$$

(see Proposition 5.1 in [22]). This along with (49) implies that

$$\|x_\epsilon^*(t_l) - x(t_l)\| \leq \sigma_1(\epsilon, T) + \sigma_2(\epsilon, T), \quad l = 0, 1, \dots, \lfloor \frac{T}{\Delta(\epsilon)} \rfloor,$$

which, in turn, leads to the estimate

$$\max_{t \in [0, T]} \|x_\epsilon^*(t) - x(t)\| \leq \sigma_3(\epsilon, T), \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \sigma_3(\epsilon, T) = 0. \quad (51)$$

From (51) it follows (see (39) and (46)) that

$$\left| \int_0^T e^{-ct} r(u_\epsilon^*(t), y_\epsilon^*(t)) dt - \int_0^T e^{-ct} \tilde{r}^*(z^*(t)) dt \right| \leq \sigma_3(\epsilon, T). \quad (52)$$

Using an argument similar to that used in Corollary 1, one can verify that (52) implies.

$$\left| \int_0^\infty e^{-ct} r(u_\epsilon^*(t), y_\epsilon^*(t)) dt - \int_0^\infty e^{-ct} \tilde{r}^*(z^*(t)) dt \right| \leq \sigma_4(\epsilon), \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \sigma_4(\epsilon) = 0. \quad (53)$$

By definition

$$\int_0^\infty e^{-ct} \tilde{r}^*(z^*(t)) dt = \tilde{R}^*(y_0) \quad (54)$$

if $u^*(y, z)$ generates an optimal ACG family and

$$\int_0^\infty e^{-ct} \tilde{r}^*(z^*(t)) dt \leq \tilde{R}^*(y_0) + \alpha \quad (55)$$

if $u^*(y, z)$ generates an α -near optimal ACG family (see Definitions 2 and 3). These along with (53) prove the proposition. \square

Proof of Proposition 11. Let $x_\epsilon^*(t) \stackrel{\text{def}}{=} (y_\epsilon^*(t), \theta_\epsilon^*(t))$ be the solution of the equation

$$\frac{dx_\epsilon^*(t)}{dt} = X(t, u^*(y_\epsilon^*(t), F(y_\epsilon^*(t))), y_\epsilon^*(t)), \quad x_\epsilon^*(0) = (z_0, 0), \quad (56)$$

where $X(\cdot)$ is as in (1). Note that a part of (56) is the equation

$$\frac{d\theta_\epsilon^*(t)}{dt} = e^{-ct} r(u^*(y_\epsilon^*(t), F(y_\epsilon^*(t))), y_\epsilon^*(t)), \quad \theta_\epsilon^*(0) = 0. \quad (57)$$

From (56) it follows that

$$\begin{aligned} x_\epsilon^*(t_{l+1}) &= x_\epsilon^*(t_l) + \int_{t_l}^{t_{l+1}} X(t, u^*(y_\epsilon^*(t), F(y_\epsilon^*(t))), y_\epsilon^*(t)) dt \\ &= x_\epsilon^*(t_l) + \epsilon \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(\epsilon\tau, u^*(y_\epsilon^*(\epsilon\tau), F(y_\epsilon^*(\epsilon\tau))), y_\epsilon^*(\epsilon\tau)) d\tau. \end{aligned} \quad (58)$$

Let $y_{y_\epsilon^*(t_l)}(\tau)$ stand for the solution of the system (1) considered on the interval $[\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]$ with the initial condition $y_{y_\epsilon^*(t_l)}(\frac{t_l}{\epsilon}) = y_\epsilon^*(t_l)$. Then, due to (41) and due to continuity of the function $X(t, u^*(y, F(y)), y)$,

$$\begin{aligned} & \left\| \frac{\epsilon}{\Delta(\epsilon)} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(\epsilon\tau, u^*(y_\epsilon^*(\epsilon\tau), F(y_\epsilon^*(\epsilon\tau))), y_\epsilon^*(\epsilon\tau)) d\tau \right. \\ & \left. - \frac{\epsilon}{\Delta(\epsilon)} \int_{\frac{t_l}{\epsilon}}^{\frac{t_{l+1}}{\epsilon}} X(t_l, u^*(y_{y_\epsilon^*(t_l)}(\tau), F(y_{y_\epsilon^*(t_l)}(\tau))), y_{y_\epsilon^*(t_l)}(\tau)) d\tau \right\| \leq \kappa(\epsilon), \quad \lim_{\epsilon \rightarrow 0} \kappa(\epsilon) = 0 \end{aligned} \quad (59)$$

(note that $F(y_{y_\epsilon^*(t_l)}(\tau)) = F(y_{y_\epsilon^*(t_l)}(\frac{t_l}{\epsilon})) \stackrel{\text{def}}{=} z_\epsilon^*(t_l) \quad \forall \tau \in [\frac{t_l}{\epsilon}, \frac{t_{l+1}}{\epsilon}]$). Since the estimate similar to (43) is valid, one can repeat the argument used in the proof of Proposition 10 to establish that the estimates (50) are valid, where $x_\epsilon(l)$, $l = 0, 1, \dots$, are defined by (48). Continuing further as in the proof of Proposition 10, one can establish the validity of (51) and, subsequently, the validity of (53). The latter along with (54) and (55) prove the proposition. \square

References

- [1] V. Arnold, V.V Kozlov, and A.I. Neistadt, “Mathematical Aspects of Classical and Celestial Mechanics, Springer-Verlag, New York, 1993.
- [2] O. Alvarez and M. Bardi, “Viscosity Solutions Methods for Singular Perturbations in Deterministic and Stochastic Control”, *SIAM J. Control and Optimization*, 40 (2001): 4, pp.1159-1188.
- [3] Z. Artstein, “Invariant Measures of Differential Inclusions Applied to Singular Perturbations”, *J. Differential Equations*, 152 (1999), pp. 289-307.
- [4] Z. Artstein, “Invariant Measures and Their Projections in Nonautonomous Dynamical Systems”, *Stochastics and Dynamics*, 4 (2004): 3, pp 439-459.
- [5] Z. Artstein and V. Gaitsgory, “Tracking Fast Trajectories Along a Slow Dynamics: A Singular Perturbations Approach”, *SIAM J. Control and Optimization*, 35 (1997): 5, pp. 1487-1507.
- [6] Z. Artstein and V. Gaitsgory, “The value function of singularly perturbed control systems”, *Applied Mathematics and Optimization* 41(2000): 3, pp. 425-445
- [7] Z. Artstein, I.G. Kevrekidis, M. Slemrod and E.S. Titi, “Slow Observables of Singularly Perturbed Differential Equations”, 20 (2007), No 11, pp 2463-2481
- [8] J.-P.Aubin, “Viability Theory”, Birkhauser, Boston, 1991.
- [9] J.-P. Aubin and H. Frankowska, “Set-Valued Analysis”, Birkhauser, Boston, 1990.
- [10] M. Bardi and I. Capuzzo-Dolcetta, “Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations”, Birkhauser, Boston, 1997.
- [11] A. Bensoussan, “Perturbation Methods in Optimal Control”, John Wiley, New York, 1989.
- [12] V.S. Borkar and V. Gaitsgory, “On Averaging of Singularly Perturbed Controlled Stochastic Differential Equations”, *Applied Mathematics and Optimization*, 56 (2007), pp. 169-209.
- [13] M.G. Dmitriev and G.A. Kurina, “Singular Perturbations in Control Problems (Review)”, *Automation and Remote Control*, 67 (2006): 1, pp. 1-43
- [14] T.D. Donchev and A.L. Dontchev, “Singular Perturbations in Infinite-Dimensional Control Systems”, *SIAM J. Control and Optimization*, 42 (2003), pp. 1795-1812.
- [15] T. Dontchev and I. Slavov, “Averaging Method for One Sided Lipschitz Differential Inclusions”, *SIAM J. Control and Optimization*, 37 (1999), pp 1600-1613.
- [16] Filar, J.A., Altman, E., Avrachenkov, K.E., 2002, “An asymptotic simplex method for singularly perturbed linear programs”, *Operations Research Letters*, v.30, no.5, pp.295-307.
- [17] J.A. Filar, V. Gaitsgory and A. Haurie, “Control of Singularly Perturbed Hybrid Stochastic Systems”, *IEEE Trans. on Automatic Control*, “46 (2001): 2, pp.179-190.
- [18] L. Finlay, V.Gaitsgory and I. Lebedev, “Duality In Linear Programming Problems Related to Long Run Average Problems of Optimal Control”, *SIAM J. on Control and Optimization*, 47 (2008): 4, pp. 1667-1700.

- [19] G. Froyland, G. A. Gottwald, and A. Hammerlindl, "A computational method to extract macroscopic variables and their dynamics in multiscale systems", *SIAM J. on Applied Dynamical Systems*, 13(2014), No 4, pp. 1816-1846
- [20] V. Gaitsgory, "Use of the Averaging Method in Control Problems", *Differential Equations* (translated from Russian), 22 (1986), pp. 1290-1299.
- [21] V. Gaitsgory. " ", Nauka, Moscow, 1991 (in Russian).
- [22] V. Gaitsgory, "Suboptimization of Singularly Perturbed Control Problems", *SIAM J. Control and Optimization*, 30 (1992): 5, pp. 1228 - 1240.
- [23] V. Gaitsgory, "On Representation of the Limit Occupational Measures Set of a Control Systems with Applications to Singularly Perturbed Control Systems", *SIAM J. Control and Optimization*, 43 (2004): 1, pp 325-340
- [24] V. Gaitsgory and A. Leizarowitz, "Limit Occupational Measures Set for a Control System and Averaging of Singularly Perturbed Control Systems", *J. Math. Anal. and Appl.*, 233 (1999), pp. 461-475.
- [25] V. Gaitsgory and L. Manic, "Use of Approximations of Hamilton-Jacobi-Bellman Inequality for Solving Periodic Optimization Problems", in *Optimization and Control Techniques*, Springer Proceedings in Mathematics and Statistics, H. Xu, K. L. Teo and Yi Zhang (Eds), 86 (2014), pp. 91-114.
- [26] V. Gaitsgory, S. Rossomakhine and L. Manic, "On Average control generating families for singularly perturbed optimal control problems with long run average optimality criteria", *Set-Valued and Variational Analysis*, 23 (2015), pp. 87-131.
- [27] V. Gaitsgory and M. T. Nguyen, "Multiscale Singularly Perturbed Control Systems: Limit Occupational Measures Sets and Averaging", *SIAM J. Control and Optimization*, 41 (2002): 3, pp. 954-974.
- [28] V. Gaitsgory and M. Quincampoix, "Linear Programming Approach to Deterministic Infinite Horizon Optimal Control Problems with Discounting", *SIAM J. of Control and Optimization*, 48 (2009): 4, pp. 2480-2512.
- [29] V. Gaitsgory and M. Quincampoix, "On sets of occupational measures generated by a deterministic control system on an infinite time horizon", *Nonlinear Analysis Series A: Theory, Methods & Applications*, 88 (2013), pp 27-41.
- [30] V. Gaitsgory and S. Rossomakhine "Averaging and Linear Programming in Some Singularly Perturbed Problems of Optimal Control", *Journal of Applied Mathematics and Optimization*, 71 (2015), pp. 195-276.
- [31] D. Givon, R. Kupferman and A. Stuart, "Extracting Macroscopic Dynamics: Model Problems and Algorithms", *Nonlinearity*, Vol. 17 (2004), R55-R127.
- [32] G. Grammel, "Averaging of Singularly Perturbed Systems", *Nonlinear Analysis*, 28 (1997), 1851-1865.
- [33] Y. Kabanov and S. Pergamenschikov, "Two-Scale Stochastic Systems", Springer Verlag, Berlin-Heidelberg, 2003.
- [34] R. Z. Khasminskii and F. C. Klebaner, "Long Term Behaviour of Solutions of Lotka-Volterra System Under Small Random Perturbations", *The Annals of Applied Probability*, Vol. 11 (2001), No. 3, pp. 952-963.

- [35] P.V.Kokotovic, H.K.Khalil and J.O'Reilly, "Singular Perturbation Methods in Control: Analysis and Design", 2nd Edition, SIAM Classics in Applied Mathematics 25, 1999.
- [36] H.J.Kushner, "Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems", Birkhauser, Boston, 1990.
- [37] A. Leizarowitz, "Order Reduction is Invalid for Singularly Perturbed Control Problems with Vector Fast Variables", *Math. Control Signals and Systems*, 15 (2002), pp. 101-119.
- [38] J.G. Llavona, "Approximation of Continuously Differentiable Functions", Mathematics Studies, 130, North Holland, Amsterdam, 1986.
- [39] S.D. Naidu, "Singular Perturbations and Time Scales in Control Theory and Applications: An Overview" , *Dynamics of Continuous Discrete and Impulsive Systems, Series B: Applications and Algorithms*, 9 (2002), pp. 233-278.
- [40] R.E. O'Malley, Jr. "Singular Perturbations and Optimal Control", In Mathematical Control Theory, W.A. Copel, ed., Lecture Notes in Mathematics, 680, Springer-Verlag, Berlin 1978.
- [41] G.A. Pavliotis and A.M. Stuart, "Multiscale Methods: Averaging and Homogenization", Springer, New York, 2008.
- [42] A.A. Pervozvanskii, V.G. Gaitsgori, 1988, *Theory of suboptimal decisions: Decomposition and aggregation*, Kluwer, Dordrecht.
- [43] V.A. Plotnikov, "Differential Equations with Multivalued Right-Hand Sides: Asymptotic Methods", AstroPrint, Odessa, 1999 (in Russian).
- [44] M. Quincampoix and J. Renault, "On the existence of a limit value in some non expansive optimal control problems", *SIAM J. Control and Optimization*, 49 (2011):5, pp. 2118-2132.
- [45] M. Quincampoix and H. Sedrakyan "Averaging problem for weakly coupled nonexpansive control systems", *Nonlinear Analysis TMA* 113 (2015), pp. 1471-158.
- [46] M. Quincampoix and F. Watbled, "Averaging method for discontinuous Mayer's problem of singularly perturbed control systems", *Nonlinear Analysis: Theory, Methods & Applications*, 54(2003), pp 819-837.
- [47] J.A. Sanders and F. Verhulst, "Averaging Methods in Nonlinear Dynamical Systems, Vol 59 of Applied Mathematical Sciences, Springer-Verlag, New York, 1985
- [48] O.S. Serea, "Characterization of the optimal trajectories for the averaged dynamics associated to singularly perturbed control systems", *Journal of Differential Equations*, 255(2013), pp. 4226-4243.
- [49] Sze-Bi Hsu, "A Remark on the Period of the Periodic Solution in the Lotka-Volterra System", *Journal of Mathematical Analysis and Applications*, 95 (1983), pp. 428-436.
- [50] V.Veliov, "A generalization of Tichonov Theorem for Singularly Perturbed Differential Inclusions", *Journal of Dynamical and Control Systems*, 3 (1997), pp 1-28
- [51] Y. Zhang, D.S. Naidu, C. Cai and Y. Zou, "Singular Perturbations and Time Scales in Control Theory and Applications: An Overview 2002-2012, *International Journal of Information Systems Sciences (IJISS)*, 9 (2014):1, pp. 1-36.

- [52] G.G. Yin and Q. Zhang, “Continuous-Time Markov Chains and Applications. A Singular Perturbation Approach”, Second Edition, Springer-Verlag, New York 2013.